An Interop Publication

ConneXions®

The Interoperability Report

March 1996 | *Special Issue: Deploying ATM Networks* | Volume 10, No. 3

ConneXions — The Interoperability Report tracks current and emerging standards and technologies within the computer and communications industry.

In this issue:

ConneXions is published monthly by Interop Company, a division of SOFTBANK Exposition and Conference Company, 303 Vintage Park Drive, Foster City, California, 94404–1138, USA.
Phone: +1 (415) 578-6900
Fax: +1 (415) 525-0194
E-mail: `connexions@interop.com`

Subscription hotline: 1-800-575-5717
or +1 610-892-1959

ISSN 0894–5926

From the Editor

Welcome to NetWorld+Interop 96 Las Vegas and to a special edition of *ConneXions—The Interoperability Report.* You are holding the one and only official technical journal of NetWorld+Interop. It has been published since the first Interop event in 1987 and covers all aspects of computer networking and interoperability.

Asynchronous Transfer Mode (ATM) is without question the hottest technology of the 1990s. At NetWorld+Interop you will find not only a number of exhibitors with ATM products, but also several conference and tutorial sessions devoted to this topic. Skeptics will tell you that ATM has been over-sold and is "just so much marketing hype," while others believe that the future of networking has its foundation firmly planted in ATM. In this edition of *ConneXions,* we look at the state of ATM from a deployment perspective. We examine a number of ATM projects in California and learn what applications are being run over ATM testbed networks.

Our first article is an in-depth look at BAGNet, the *Bay Area Gigabit Network.* Fifteen San Francisco Bay area organizations have been participants in BAGNet funded by Pacific*Bell under a California Research and Educational Network (CalREN) grant. The article provides a review of BAGNet and its goals, gives a brief overview of ATM, and details experiences with BAGNet and its use.

The *Monterey BayNet ATM Project* is another of the CalREN funded testbeds. The network uses ATM as an enabling technology for tele-education, tele-science, and electronic libraries linking the Silicon Valley, Santa Cruz, and The Monterey Peninsula. Our second article looks at BayNet applications, infrastructure, and participation.

ATM offers a number of new capabilities. But "pure" ATM networks are still quite far away. Systems that take advantage of ATM will continue to be directly attached to traditional media such as Ethernet and will make use of existing network layer protocols such as IP. This means that in order to effectively use ATM, there must be efficient methods available for operating multiple internetwork layer protocols over heterogeneous networks made up of ATM switches, routers, frame switches, and hubs. Our final article discusses routing options in the "multiprotocol over ATM" environment.

If you find yourself struggling with a particular aspect of networking we might have the answer for you in one of our 108 back issues. For a complete index, see `http://www.interop.com` or send e-mail to `connexions@interop.com` with your request. We can either send you the index file electronically or in hardcopy. Finally, if you're not already a *ConneXions* subscriber, take advantage of the special conference discount and sign up today.

BAGNet

Experiences with an ATM Metropolitan-area Network

by

**Dave Wiltzius, Lawrence Livermore National Laboratory,
Lance Berc, Digital Equipment Corporation,
Siddhartha Devadhar, Pacific*Bell**

Introduction

Fifteen San Francisco Bay area organizations have been participants in BAGNet, an IP over ATM (Asynchronous Transfer Mode) testbed. BAGNet has been funded by Pacific*Bell under a CalREN (California Research and Educational Network) grant for about 2 years. This ATM network proved to be a somewhat ambitious project given the maturity level of ATM technologies at the time, but ATM networks of this scale should be routine within a few years.

We will provide a review of BAGNet and its goals, give a brief overview of ATM, then detail our experiences with BAGNet and how we have used this network. We have also solicited information from participants of BAGNet and other CalREN testbeds on their experiences, activities, and the applications deployed. URLs are provided at the end of this article for additional Web based information pertaining to BAGNet and other CalREN testbeds.

The content of this article, particularly the opinions and observations, belong to the authors and in no way represent the BAGNet community, or any organization or other members of the BAGNet community. Finally, any opinions or observations in this article are those of the authors, not of their employers.

BAGNet

BAGNet (*Bay Area Gigabit Network*), one of several CalREN grants funded by Pacific*Bell, began in early 1994. The BAGNet participants consist of fifteen commercial, educational, and research organizations in the San Francisco Bay Area: Apple, Digital Equipment Corporation (DEC, Digital), Hewlett-Packard (HP), International Computer Science Institute (ICSI), Lawrence Berkeley National Lab (LBNL), Lawrence Livermore National Lab (LLNL), NASA-Ames, Pacific*Bell (Pac*Bell), Sandia National Laboratories/CA (SNL), Silicon Graphics Inc. (SGI), SRI International (SRI), Stanford University, Sun Microsystems (SUN), University of California at Berkeley (UCB), and Xerox Palo Alto Research Center (Xerox-PARC). These fifteen participants are connected by a high-speed (155 Mb/s) ATM metropolitan area network provided by Pac*Bell under the CalREN grant. Each site is minimally required to support Classical IP over ATM according to RFC 1577 [1]. BAGNet offers high-performance communications using the IP protocol suite, featuring high-speed motion JPEG compressed video via IP multicast.

Purpose

The purpose of the CalREN grants such as BAGNet is to develop, deploy, and demonstrate applications that are enabled by a high-performance, long-distance communication infrastructure. A tele-seminar application is the showcase application of BAGNet. Several BAGNet participants offer live and recorded seminars and courses to the BAGNet community with high quality video (30 frames per second, 320×240 resolution) and audio. BAGNet is also used for a variety of other applications, several of which will be described below.

Technology comparison

Each BAGNet site has workstations that can communicate over BAGNet at up to 155Mb/s. In comparison, typical local area network (LAN) technologies use Ethernet (10Mb/s), FDDI (Fiber Distributed Data Interface, 100Mb/s) and Fast Ethernet (100Mb/s).

Typical wide area and metropolitan area networks utilize Frame Relay (0.64–45Mb/s [2]), SMDS (Switched Multimegabit Data Service, up to 45Mb/s) and ISDN (Integrated Service and Digital Network, 0.064Mb/s and 0.128Mb/s). Data communication over regular telephone lines from a home computer with a modem is often less than 19,200 baud or about 0.020Mb/s.

BAGNet offers a lot of bandwidth to the desktop over long distances, generally more bandwidth than available to a modern workstation on a LAN.

ATM Overview

Many readers are probably well versed in ATM and the associated technologies. However, to make the information that we are providing on BAGNet more relevant to a wider audience we will first present a brief overview of ATM.

BAGNet is based on ATM, a communication technology that the telecommunication industry began developing in the mid-1980s. ATM was designed for the deployment of BISDN (Broadband Integrated Services and Digital Network), anticipating the need to provide voice, (high-resolution) video, and (high-performance) data over the same infrastructure to the office and home. With the present generation of high-performance, inexpensive multimedia computers and applications such as the MBone (*Multicast backBone*) tools, the desktop computer is able to assimilate and present many forms of information.

BISDN was advanced by the telecommunication industry with several features in mind [3]:

- *Flexibility:* high and multirate bandwidth,
- *Capability:* e.g., voice, video, high-speed data services,
- *Versatility:* unified transport network supporting the above services.

The layers

The vision offered by BISDN takes advantage of several technologies also utilized by BAGNet:

- *Fiber optics:* The telecommunications infrastructure is rapidly being replaced by fiber optics which offers many advantages over copper and wireless transmission systems: higher bandwidth, higher reliability, longer repeater spacing, greater security, smaller and lighter components, greater growth potential, and lower system costs [2].
- *SONET:* (Synchronous Optical NETwork). This protocol provides for a flexible yet unified means to provision bandwidth between two components. The provisioning of this bandwidth is relatively static and under control of the service provider (Pac*Bell in the instance of BAGNet). In an appropriately architected network, SONET can enhance the fault tolerance of the provisioned paths by automatically and quickly (within 50 milliseconds [2]) restoring service upon detection of errors (e.g., fiber damage from construction, high errors due to component failure).

  SONET overhead is approximately 3 percent of the bandwidth (e.g., a SONET payload of 149.76Mb/s for a 155.52Mb/s OC-3c provisioned path).

  In BAGNet each site is provisioned an OC-3c SONET path to one of Pac*Bell's BAGNet backbone switches in Oakland or Palo Alto.

continued on next page

BAGNet Experiences *(continued)*

- *ATM:* The SONET components provide a path between the BAGNet ATM switches and the site ATM switches. ATM provides a virtual connection or circuit between two end-systems (hosts, routers, etc.). The virtual circuit is defined by virtual channels (VC), which only have local significance in the ATM switches and end-systems.

  ATM overhead is approximately 9.5 percent of the bandwidth, hence the 155Mb/s network has an ATM payload capacity of about 135Mb/s (the SONET overhead must also be deducted).

  To provide IP over ATM there is an additional layer. Each ATM cell is 53 bytes consisting of a 5 byte header and a 48 byte payload. This was a compromise between the ANSI proposal of a 5 byte header and 64 byte payload, and the European Telecommunications Standard Institute proposal of a 4 byte header and 32 byte payload [4, 5]. Although this is adequate for a quantum of digitized audio it is too small for IP. Classical IP over ATM specifies IP packets be encapsulated over ATM using AAL5 (ATM Adaptation Layer 5) with a default IP MTU (*Maximum Transmission Unit*) of 9180 bytes. AAL5 is just one of several ATM adaptation layers defined by the ATM standards.

Concepts for BAGNet

The following concepts were important in the successful deployment and utilization of BAGNet, and they will provide the foundation for the discussion of BAGNet that follows.

- *PVCs, SVCs and VPs* [6, 7, 8]: The ATM header is 5 bytes consisting of several fields, two of which are the *Virtual Path Identifier* (VPI) and the *VC Identifier*(VCI). The VPI and VCI (VPI/VCI) are set by the transmitting host adapter (i.e., network interface) to specify the destination of the ATM cells. When receiving ATM cells, the host adapter uses the VPI/VCI to demultiplex the cells for reassembly to IP packets.

  ATM switches use the VPI/VCI to determine how to route the ATM cell. If this is a *Permanent Virtual Circuit* (PVC), then the ATM switches and ATM hosts have been pre-configured with the routing information. This is the case with BAGNet. If *Switched Virtual Circuits* (SVCs) are used, then a host interacts with the ATM switch via "signaling" to dynamically establish a virtual circuit between end-systems. Using PVCs, BAGNet had to configure a PVC between each host in BAGNet and every other host in BAGNet—a PVC mesh.

  Virtual paths (VPs) can be viewed as a bundle of virtual circuits with an origination and termination end point. VPs can be used to configure routing information in ATM switches which can be transparently used by virtual circuits.

- *Segmentation and Reassembly (SAR):* With Classical IP over ATM, an IP packet is encapsulated using AAL5. For an IP packet to be received it is necessary to remove and process the AAL5 information and, for each ATM cell, extract the 48 byte payload for the next chunk of the IP packet. This is done within the context of the VC since a host could be receiving cells from several other hosts, each sending to this host on a unique VC. This reassembled IP packet would then be processed by the host. For an IP packet (or fragment) the size of the default MTU of 9180 bytes, this would have to be done for 192 ATM cells.

For transmission of an IP packet, the reverse occurs where the IP packet is segmented into a number of ATM cells.

With early ATM host adapters the host participated in the SAR of the ATM cells. However, the latest ATM host adapters perform SAR with minimal host intervention, and often in silicon or firmware. One issue associated with SAR is the number of simultaneous VCs it will support (i.e., how much memory is available to reassemble IP packets). Another issue is the speed with which SAR can be performed.

- *Bandwidth limited VCs:* ATM SVCs, in principle, provide bandwidth on demand where guaranteed bandwidth can be allocated as the virtual circuit is established. Presently many service providers market ATM PVCs by the Mb/s [9] (e.g., bandwidth limit), which is enforced by the service provider's ATM switches. In this case, the ATM switch can drop cells if they are received faster than the configured bandwidth limit of that PVC.

  To adhere to the bandwidth contract, the sending system should limit the transmission bandwidth. This functionality is provided by the host adapter hardware and software by scheduling the ATM cells to be transmitted (i.e., cell pacing or scheduling). In high-performance ATM host adapters this cell scheduling can be quite precise. See the comments in the "Performance Measurements" section below, where an HP ATM analyzer was used to compare this feature on several workstations for two ATM host adapter vendors.

- *Multicast:* This is the capability for the transmission from *one* source to be delivered to *many* receivers (contrast with unicast where the transmission from one source is received by one destination).

  At the ATM level, multicasting is specified by appropriately configuring multicast VCs into the ATM switches. First the source is configured (specifying the VC and ATM port on which the cells to multicast are received—termed the "root"), then one or more branches are configured (the VC and ATM ports to which the multicast cells are forwarded). The switches will then replicate the ATM cell or otherwise effect a multicast for each multicast branch. The exact method used is dependent upon the architecture of the ATM switch. The performance degradation in the ATM switch experienced for multicast traffic varies, but for switches with the appropriate architecture it can be low.

  Multicast is required at the IP level for use by the MBone application tools used on BAGNet. Most UNIX workstations currently support multicast IP over shared network media such as Ethernet and FDDI. This was not initially the case for ATM and was an issue that BAGNet had to address.

Unique aspects

In the time frame of its deployment, BAGNet offered several unique features:

- *Multicast IP over multicast ATM:* There are provisions in multicast IP to allow it to operate over non-multicast network media that utilizes the *mrouted* daemon to tunnel multicast IP traffic over TCP/IP connections. Multicast IP could also be implemented by using a multicast "server."

continued on next page

BAGNet Experiences *(continued)*

Instead, the BAGNet community chose to establish a scalable ATM architecture. On BAGNet, multicast IP uses a multicast ATM PVC mesh where each site would be provided 4 unidirectional multicast PVCs on which to transmit. Sites would receive multicast IP on this multicast mesh from the incoming PVCs.

- *Audio and video:* High-speed motion JPEG video and high-quality audio would be transmitted on multicast IP over multicast ATM.
- *The ATM environment was heterogeneous:* Sites were encouraged to use a variety of ATM hardware and software products. BAGNet consists of ATM switches from about five different ATM vendors, and ATM host adapters and software from about eight different vendors. ATM interoperability problems, apart from misconfigurations, were surprisingly few.
- *A large PVC mesh connecting a large number of organizations:* BAGNet's initial proposal was downsized due to limitations in the BAGNet backbone switches. Still, BAGNet has a unicast PVC ATM mesh connecting 60 hosts at 15 sites (although the connectivity has been verified by inspection, only about 40 hosts have actually appeared on BAGNet at this time with about 25 regularly connected). Additionally, each of the 60 hosts at the 15 sites has a multicast ATM PVC configured.

  The BAGNet community was eager to utilize SVCs for several reasons: Ease of setting up the ATM network; flexibility of the network configuration; and exposing the network researchers to additional interesting and relevant issues at the application level. Unfortunately, the BAGNet ATM backbone switches were put into service in 1993 and did not support SVCs.

  Another option for BAGNet that recently became available, but not yet explored, is to establish a mesh of permanent virtual paths (PVPs) complementing the existing PVC mesh. Pairs of sites with ATM switches that have an interoperable SVC implementation could use the PVP mesh to "tunnel" the SVC signaling information (by passing through the backbone switches, which do no understand SVCs). Also, sites could configure additional PVCs in the PVP mesh.

Architecting BAGNet

The BAGNet backbone switches are located in Pac*Bell's Central Offices (CO) in Oakland and Palo Alto, connecting the BAGNet sites in a network about 75km in diameter. Although BAGNet expected one ATM switch in each CO, Pac*Bell in fact deployed BAGNet utilizing three ATM switches in order to support other CalREN projects in the Bay Area. There are several OC-3c trunks interconnecting the BAGNet backbone ATM switches (see Figure 1 below).

BAGNet service is provided to each site as a single-mode fiber pair. Some BAGNet sites were a very short distance from the CO, while other sites were routed through SONET equipment to span the 50km fiber run to the CO.

BAGNet was deployed as a Classical IP over ATM network consisting of one Class C network. SVCs were not available in the lifetime of the project. Virtual paths were not used as it was unclear if the BAGNet backbone switches supported them, and it was certain that some vendor products (host adapters and site ATM switches) used on BAGNet did not support them. So PVCs were used instead.

Due to the limited number of VCs supported per port on the BAGNet backbone switches, each of the 15 BAGNet sites was allocated 4 unicast IP addresses. Additionally, the multicast PVC mesh provided for 4 multicast PVCs per site.

There were a few individuals in the BAGNet community that had experience configuring ATM networks using PVCs. Two of these individuals, Mark Laubach then with HP, and Berry Kercheval of Xerox-PARC, were the BAGNet architects. They provided the IP and PVC databases for the sites, and provided the PVC configuration information to Pac*Bell for the BAGNet backbone switches.

PVC configuration

The following provides some idea of the ATM configuration that was required to get BAGNet working. BAGNet consists of a unicast PVC mesh which is bidirectional, and a multicast PVC mesh which is unidirectional.

To send data, the IP code must translate from the destination IP address to a media address. For certain kinds of network media (broadcast capable) this is dynamically performed using the *Address Resolution Protocol* (ARP). The IP address to media address mapping is established by discovering hosts using broadcasts and is maintained in an ARP cache. For ATM with PVCs, the host is pre-configured with static ATM ARP cache entries to map from an IP address to a PVC.

At the time BAGNet began, the standards for Classical IP over ATM were nearing completion and ATM host adapter products did not have a Classical IP over ATM configuration typical of current products. Instead the ATM host adapter software had to be carefully configured to effect Classical IP over ATM. Each incoming and outgoing PVC was configured for AAL5 and the appropriate link-level encapsulation method. This information was provided in the ATM ARP cache configuration for each PVC on each BAGNet host.

For each site, the unicast PVC configuration is quite manageable. There are 60 hosts on the BAGNet, so the ATM ARP cache had 60 unicast entries (a loopback was provided for each host to the BAGNet backbone). This ATM ARP cache configuration was identical for all BAGNet hosts at all sites.

For each site's ATM switch, it was necessary to create virtual circuits from each of the 4 site hosts to all of the 60 BAGNet hosts. This is a total of 480 VCs where the doubling occurs because VCs are unidirectional and the PVC mesh is to be bidirectional.

The Pac*Bell BAGNet backbone switches must configure the mesh that interconnects the site's ATM switches. Not counting the loopback virtual circuit for each host, the total number of virtual circuits for the BAGNet backbone is (60 * 59) / 2 (the sum of the connections for each site: 59+58+...+1) = 1770. Double this to get bidirectional connectivity and include the 60 loopback PVCs, and Pac*Bell had to configure 3,600 virtual circuits for the unicast PVC mesh!

In comparison, a unicast mesh using virtual paths would require only one path from each site to all other sites, or (15 * 14) / 2 = 105, which is 310 virtual paths for a bidirectional VP mesh.

The unidirectional multicast PVC mesh is easier. Each site was allocated 4 PVCs for multicast ATM transmissions so each host would have one entry in the ATM ARP cache for transmitting multicast IP traffic.

continued on next page

BAGNet Experiences *(continued)*

Each host was also configured to receive the multicast transmissions from the other 59 BAGNet hosts, the purpose being to specify AAL5 and the link-level encapsulation method for each incoming multicast PVC. So each host had 60 ATM ARP cache entries for the multicast PVCs.

The site ATM switch was configured to multicast each of the other site's incoming multicast PVCs, minimally to each BAGNet host at the site. This required the creation of 56 multicast roots (one for each of the 4 hosts at the other 14 sites), each multicast root having 4 branches to each of the site's hosts for a total of 224 VCs. The site's 4 outgoing multicast PVCs required an additional 4 unidirectional VCs.

To support ATM multicast, Pac*Bell had to configure the BAGNet backbone to distribute 4 multicast PVCs from each site to the other 14 sites. For one site this would be 56 virtual circuits and for all 15 sites this is 840 virtual circuits.

Hence a fully configured BAGNet site requires 120 ATM ARP cache entries per host, and 704 VCs on the ATM switch. For the BAGNet backbone, Pac*Bell had to configure about 4,500 virtual circuits traversing from 1 to 3 backbone ATM switches.

Although these numbers appear onerous, the PVC configurations were automatically generated. The BAGNet community generated and distributed scripts to configure the ATM host adapters and ATM switches. Since nearly all ATM products supported this script driven capability, generally the site configuration was easy and error-free. The challenge for the sites was in dealing with the few peculiarities that some of the ATM products exhibited (e.g., limited number of VCs supported on a host adapter, default maximum VCI of a switch, etc.).

Unfortunately, using scripts to configure the BAGNet backbone ATM switches was not possible. We do not fully understand the reasons, but one contributing factor was the fact that the ATM switches were in a Pac*Bell Central Office, and BAGNet had to abide by the CO operational policies. Hence the PVC configuration was established by Pac*Bell's Network Operation Center using other means, which were less automated than scripts.

Teleseminar application

The teleseminar application used by BAGNet was initially developed for use in the Internet and is based on the MBone tools. For teleseminars, BAGNet sites generally used *sd* (session directory), *vic* (video), and *vat* (audio). These tools allow the advertising and transmission of teleseminar sessions over UDP/IP. These teleseminar sessions can be received by one or many other computer workstations that opt to join the session. While the workstation that is the source of an MBone session requires multimedia hardware (e.g., microphone, camera, video card), the workstation receiving the session and opting for software video decompression only requires the MBone tools that are freely distributed on the Internet for many types of workstations.

Multimedia sessions can function as a conference (one speaker, the rest are part of the audience and interact relatively occasionally with the speaker such as for questions) or meetings (all session participants can contribute equally to the session) in a coordinated manner through the use of the MBone tools. For teleseminars, only one workstation at a site would need to be capable of providing high-speed video and audio to BAGNet from the site's conference room.

BAGNet was to offer the teleseminar capability with very high-performance video and audio. Since it was desired that each site be capable of providing teleseminar content, reasonably priced commercial video compression hardware available for several workstations was desired. Hence motion JPEG was chosen as the video compression format for the teleseminars. The MBone video tool *vic* was developed by Steve McCanne of LBNL and enhanced to include motion JPEG with software decompression, and support for several motion JPEG compression boards.

Problems and issues encountered

Deploying BAGNet presented a variety of problems, not all technical, that were addressed with varying degrees of success. Members of the BAGNet community worked together to resolve issues and exchange information on the BAGNet mailing list and the Web. In general, we found many opportunities to acquire invaluable experience and knowledge pertaining to ATM networking. These experiences are shared below, and we hope that our candid comments will prove helpful to others.

The first issue many sites encountered was in dealing with the single-mode fiber for the connection to the BAGNet backbone switches. In many cases the initial issue was availability of single-mode ATM ports or host adapters, followed by cost. Single-mode optics are much more expensive than multi-mode optics (about $5K per single-mode port, versus $1.5K per multi-mode port in 1994), and ATM switch modules often included 4 ports. Hence sites generally used one of several options for connecting to BAGNet (in order of decreasing cost): 1) acquire a host adapter with single-mode optics and connect just one workstation, 2) convince a switch vendor to modify one multi-mode port to a single-mode port, 3) purchase a multi-mode to single-mode converter, 4) purchase a single-mode module for the site ATM switch.

Single-mode fiber is used to span long distances (about 50km for SONET OC-12 [2]) between components or repeaters. The optical receiver for a single-mode fiber will saturate when the signal is too high, which occurs when the distance between components is short. This was the scenario for some BAGNet sites, and was unexpected and unanticipated. The solution is to introduce attenuators (quite inexpensive) in the fiber path.

BAGNet deployed IP over ATM as per RFC 1577, "Classical IP over ATM." Once a few sites provided hosts that were correctly configured for IP over ATM, it became easy for other sites to verify ATM configurations of their hosts. However, the multicast PVC mesh to support multicast IP presented a somewhat more difficult problem since all other hosts would receive the multicast ATM transmissions. This was aggravated by the tendency of some host adapter software to crash the receiving host when the transmitting hosts were misconfigured (e.g., wrong link-level encapsulation). Hence most BAGNet sites would not configure an incoming multicast PVC until the transmitting host was first validated by some other site.

By mid-1994 all BAGNet sites had hosts connected and it was discovered that the PVC mesh connectivity was incomplete. This was puzzling as many sites had rather elaborate local ATM networks using PVCs and were experiencing no connectivity problems with those networks. It was unknown if we were dealing with configuration problems, host connection issues (some sites would temporarily disconnect a host from BAGNet), ATM interoperability problems, or reliability problems in the ATM hardware and software products.

continued on next page

BAGNet Experiences *(continued)*

With the help of the Pac*Bell Network Operations Center (NOC) it was ascertained that most connectivity problems were in the configuration of the BAGNet backbone switches (recall that these backbone switches were configured manually). The BAGNet community then embarked on developing connectivity monitoring tools. Lance Berc at Digital wrote the versatile *bagping* utility which performed the IP *pings,* formatted the results into a Web "ping page" and e-mailed the connectivity information to a configurable address, all at some specified interval (e.g., hourly). Many sites used bagping to update their ping pages and also e-mail their connectivity to Apple where it was compiled into a connectivity matrix published on the Web (thanks to Alagu Periyannan at Apple). With this information in hand, it was determined that the problem was a lack of robustness in the commercial configuration tools used to configure the BAGNet backbone.

Support for multicast ATM was generally available in ATM switches. However there was poor support in the host adapter software for multicast IP over multicast ATM (only DEC offered this feature in 1994). The problem was the inability to establish an ATM ARP cache entry that would map a subset of multicast IP addresses to a PVC for transmission. Contrast this with a unicast IP ATM ARP cache entry, which would map a single IP address to a VC. A workaround was devised: Configure a single multicast IP address to the multicast ATM PVC. This worked for host adapter software from many (but not all) vendors. Incoming multicast PVCs was not a problem in general, and could be configured just like the incoming unicast PVCs.

Soon after ping pages were appearing Lance Berc tried some *pings* with larger packets (the default *ping* packet was 56 bytes). It was noticed that the *pings* with larger packets would rarely succeed to hosts at sites connected through at least two BAGNet switches. Further investigation by Lance and others provided enough information so that Pac*Bell was able to determine that the problem was due to a misconfiguration of the BAGNet backbone PVCs. Specifically, it was intended that the PVCs would be configured with no bandwidth limitation ("best effort"). Instead, the PVCs were inadvertently configured with a bandwidth limitation of 140Mb/s, which was only detectable when the ATM host adapters sent long bursts of consecutive cells—a unique capability at the time of the Digital ATM host adapters. Lance provided the analysis details on the Web.

The teleseminar application was the showcase BAGNet application. It was our goal that sites would provide teleseminars on BAGNet featuring high-performance audio/video, interesting technical content, and widely acclaimed speakers. This placed some specific requirements on the sites such as: A BAGNet host capable of transmitting high-performance audio/video, a conference room set up for quality audio/video, an audio/video feed from that conference room to the transmitting BAGNet host, and the appropriate legal waivers from the speakers. These requirements proved to be quite difficult for some sites to overcome, yet there have been several notable teleseminars successfully presented on BAGNet—some ongoing—that are discussed further below.

Once some high-performance audio/video content was available on BAGNet, it was found that few hosts could receive a 2–5Mb/s motion JPEG stream over ATM, perform the JPEG decompression in software, and display on the workstation at 30 frames per second (fps) at 320 × 240 resolution. Even some hosts with hardware JPEG decompression assistance did not fare well.

This is quite evident with the video tool *vic*. This tool initially displays the incoming video streams as a small icon at a reduced frame rate, and will also display the received fps, data rate, and loss rate. On one low-end workstation on BAGNet, a 2Mb/s motion JPEG stream was being received by *vic* with no loss when just the icon was displayed. When the video stream was then displayed at 320 × 240 resolution using software decompression, *vic* reported 5fps, 0.6Mb/s, and a loss rate of about 70%. However, high-performance workstations, or workstations with certain hardware decompression equipment, could easily receive and display a 5Mb/s, 30fps motion JPEG video stream over ATM using *vic*.

There is interest in the BAGNet community to further investigate the issues pertaining to the transmission and reception of high-performance audio/video.

BAGNet and other CalREN project activities

The CalREN funded Pac*Bell ATM testbeds have been successful in demonstrating several applications utilizing a high speed ATM connection in a metropolitan and wide area network. These ATM testbeds have been used for the participation of several conferences and other special demonstrations, as well as being used for a variety of research activities. Individuals from some Pac*Bell ATM testbed sites describe some of the activities for which they used the testbeds:

***Project:* Teleseminars**

Several sites have transmitted teleseminars that featured audio and high-performance (i.e., several Mb/s) motion JPEG video. LLNL transmitted seminars by several internationally renown scientists as part of the Director's Distinguished Lecture Series, UCB has transmitted several seminars, Stanford has transmitted two semesters of a weekly seminar course featuring entrepreneurs and experts in ATM technologies, and Apple and Xerox-PARC continue to transmit a weekly seminar. Several sites regularly transmit audio and high-performance video content on BAGNet.

In addition to other sites transmitting teleseminars over BAGNet, Sandia, LBNL, and LLNL have been transmitting CSPAN audio and video. The continuous audio and video feeds have allowed researchers at Sandia, LLNL and Xerox-PARC to test new versions of LBNL's audio and video applications and study the quality of workstation-based audio and video. The video streams are 320×240 frame-encoded JPEG with a frame rate of 20–24 fps at bandwidths of 1–2.5 Mb/s. The audio is Pulse Code Modulation (PCM) encoded with a data rate of 78 Kb/s.

***Project:* ACM Multimedia 94**

As part of a panel discussion in ACM Multimedia 94, one of the panelists participated remotely using BAGNet and the MBone tools. (Contact: Bill Johnston, LBNL, `johnston@george.lbl.gov`).

***Project:* ATM Traffic Recording**

At the high speeds of ATM, comprehensive performance simulations of even a single switch take impossibly long times. Network level performance simulations for ATM traffic are not feasible to conduct on available computing power. Real traffic measurements are necessary for traffic engineering and capacity planning. By understanding the nature of these relationships which are currently unknown, we hope to learn about the network resource allocation requirements in ATM networks. To this end a joint Pacific*Bell and Bellcore project using specially designed Bellcore hardware was initiated.

The Bellcore traffic recorder is capable of time-stamping and recording every ATM cell that flows by, up to the OC-3 rate. Bellcore proprietary hardware that performs the time-stamping is controlled by a Sun SPARCstation.

continued on next page

BAGNet Experiences *(continued)*

A Sony high speed tape recorder is controlled by a dedicated processor also running UNIX. Data transfer occurs over HIPPI and other proprietary interfaces within the system. Each cartridge can store up to 96 Gbytes of data. At OC-3 rates this translates to approximately 4–5 hours of recording time at full load. Soliciting the participation of CalREN users was problematic. Several CalREN projects use the ATM testbed network to transfer customer data which has implications for confidentiality.

BAGNet is primarily composed of universities, national labs and leading edge high technology companies that are investigating the use of high-speed broadband networks, and developing applications that can exploit these high bandwidths. BAGNet being a primarily research-oriented group of participants, there was sufficient interest in measuring the ATM traffic generated as a result of traditional applications as well as that generated by ATM-specific applications. Hence it was a natural choice for the recording effort. Between September 11, 1995 and October 6, 1995 traffic traces totaling 400 Gbytes were collected on BAGNet.

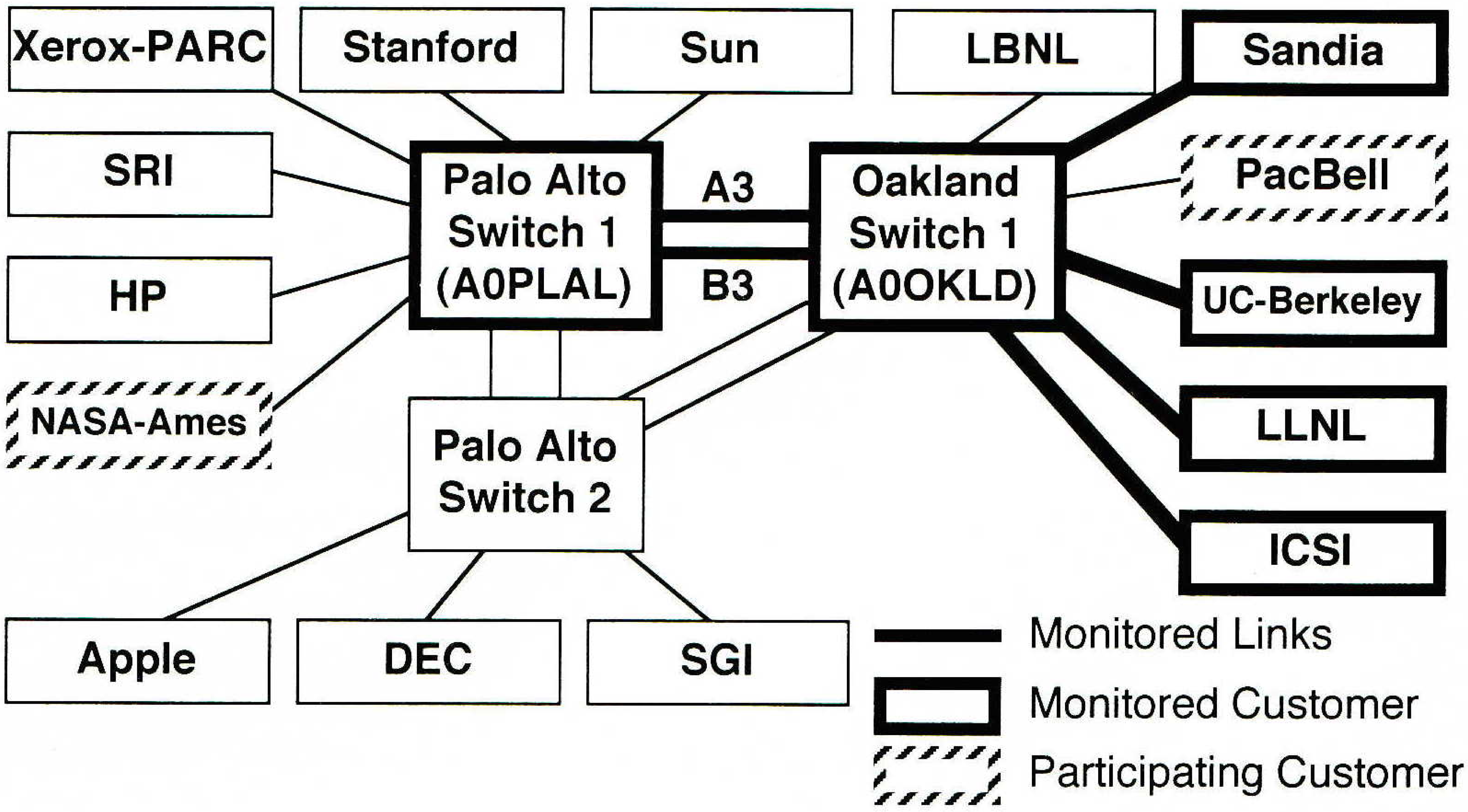

Figure 1: ATM cell recording on BAGNet

Data Collection

Optical splitters were installed on OC-3 lines in the Oakland Central Office on two trunk ports and four customer ports. The two trunks between the A0OKLD and A0PLAL switches are connected to ports "A3" and "B3" on the A0OKLD switch and are therefore referred to as the "A3 trunk" and "B3 trunk" respectively. See Figure 1 for the BAGNet topology. Starting September 11, 1995 the OC-3 lines were monitored. After a week of verification and making preliminary observations on various lines, actual data collection began on September 18, 1995. Lawrence Livermore National Lab's (LLNL) access was monitored for this duration. In the week of September 25, 1995 a coordinated measurement was performed in cooperation with BAGNet users. Finally, from September 29 until October 6, 1995 the "A3" trunk was monitored.

Analysis

Since this is the first experiment of its kind, many new tools had to be developed for post processing of the large data sets. Bellcore has developed a suite of applications for the analysis of large scale data. One of the tools (tribeca) is a database querying language that can use data streams for online manipulation of packet data.

Template structures have been defined for ATM cells, ATM AAL5 packets, IP packets and TCP messages. After an initial analysis at the ATM cell level, Bellcore will perform packet reassembly at the higher layers. Of particular interest is the nature of the timing relationships between the packets at each layer. This relationship is affected by the particular protocols used at each layer as well as the interaction between the layers.

Summary Remarks

The entire project was an excellent learning process for all participants. Live network monitoring, coordinating experiments among participants and large scale data gathering and processing has different logistic requirements. All appropriate information resulting from this study will be made available to the participants and the public. This experiment will provide invaluable data for further understanding of ATM traffic. It also paves the way for similar and more refined studies in the future that include not only ATM but other Fast Packet services. (Contact: Siddhartha Devadhar, Pacific*Bell, `sydevad@srv.pacbell.com`).

***Project:* Collaborative Science**

RIACS/NASA-Ames and Sandia are developing and evaluating an environment to support real-time collaborative scientific work, using high-speed ATM networks as the communications medium. Our environment incorporates tools developed by Sandia.

We conducted a collaborative work session over BAGNet between RIACS and Sandia during the recent Bellcore traffic recordings. We plan to analyze this traffic data to obtain an understanding of communication patterns for collaborative applications. RIACS is also working with NASA earth scientists to evaluate our collaborative environment for interactive data analysis and remote training. (Contact: Marjory Johnson, RIACS/NASA-Ames, `mjj@riacs.edu`).

***Project:* Distance Learning**

Changing occupational requirements, increasing demographic diversity in society, and growing cost pressures in our educational system are calling for a radical redesign of our learning paradigm. Asynchronous or on-demand education is therefore of critical interest. A project to investigate the educational, technological, and economic facets of distance learning, known as the *Asynchronous Distance Education ProjecT* (ADEPT), is presently being carried out at the Center for Telecommunications at Stanford in conjunction with the Stanford Center for Professional Development (SCPD). ADEPT is being funded by a grant from the Sloan Foundation.

The goal of ADEPT is to provide on-demand access to selected Stanford engineering classes to remote professionals and students on campus. The current mix of students on and off-campus is 50/50. To this end, class material is digitized for popular computer platforms and then made available on servers connected to the Internet and BAGNet. ADEPT participants may download a class or play it back in real-time directly from their desktops. During the 1995–96 academic year, twelve graduate-level engineering courses, or four per quarter, will be distributed. In autumn quarter 1995, there were a total of approximately 100 students, both on and off-campus, taking ADEPT classes for Stanford credit

The first step is to digitally capture a class' video, audio, and handouts. Video and audio are digitized simultaneously, on four different platforms: PC (Windows), Macintosh, Sun Microsystems, and Silicon Graphics. The video and audio files are then transferred by FTP from the capture stations to one of two servers: Real-time Playback (RTPB) and File-Download (FD).

continued on next page

BAGNet Experiences *(continued)*

In RTPB mode, clients receive a video and audio stream directly from the server, which they can play in real time. In FD mode, clients must download files to their local host before playback. In both RTPB and FD, clients use a World-Wide Web browser such as Netscape as the principal front-end. The ADEPT home page can be found at URL: `http://adept.stanford.edu`.

Corporate sites off-campus sometimes set up a satellite FD server, mirroring files for the ADEPT classes that students have subscribed to. Clients can then download files from the satellite server over the corporate LAN. Of course, they can also download directly from the Stanford server, assuming that the corporate firewall permits this. This distribution mode is advantageous for two reasons. First, sites enjoying BAGNet connections can rapidly FTP files from the Stanford server onto their satellite server. Second, sites can then re-distribute files over their own ATM or Ethernet LANs and thus bypass the Internet throughput bottleneck.

Lecture notes, encoded in *PostScript* and Portable Document Format (PDF), are also made available on both servers. Video and audio, however, are encoded differently for FD and RTPB modes. For RTPB, the video is compressed and encoded using MPEG and an experimental algorithm, developed at Stanford, based on Vector Quantization (VQ). For FD mode, video and audio are encoded in a variety of formats, depending on the platform. A Cell-B compressed format is encoded exclusively for Sun workstations. A QuickTime format is encoded for PCs, Macintosh computers, and UNIX workstations. The storage and throughput requirements required for a 75 minute lecture is substantial. For example, MPEG-encoded video at 640×480 pixels requires network throughput of 2 Mb/s and just over 1 GB of storage capacity. Hence, the high data transfer rates afforded by ATM are crucial for both FD and RTPB modes.

The experimental ATM network used to distribute ADEPT material over the Stanford campus, shown in Figure 2, is very heterogeneous.

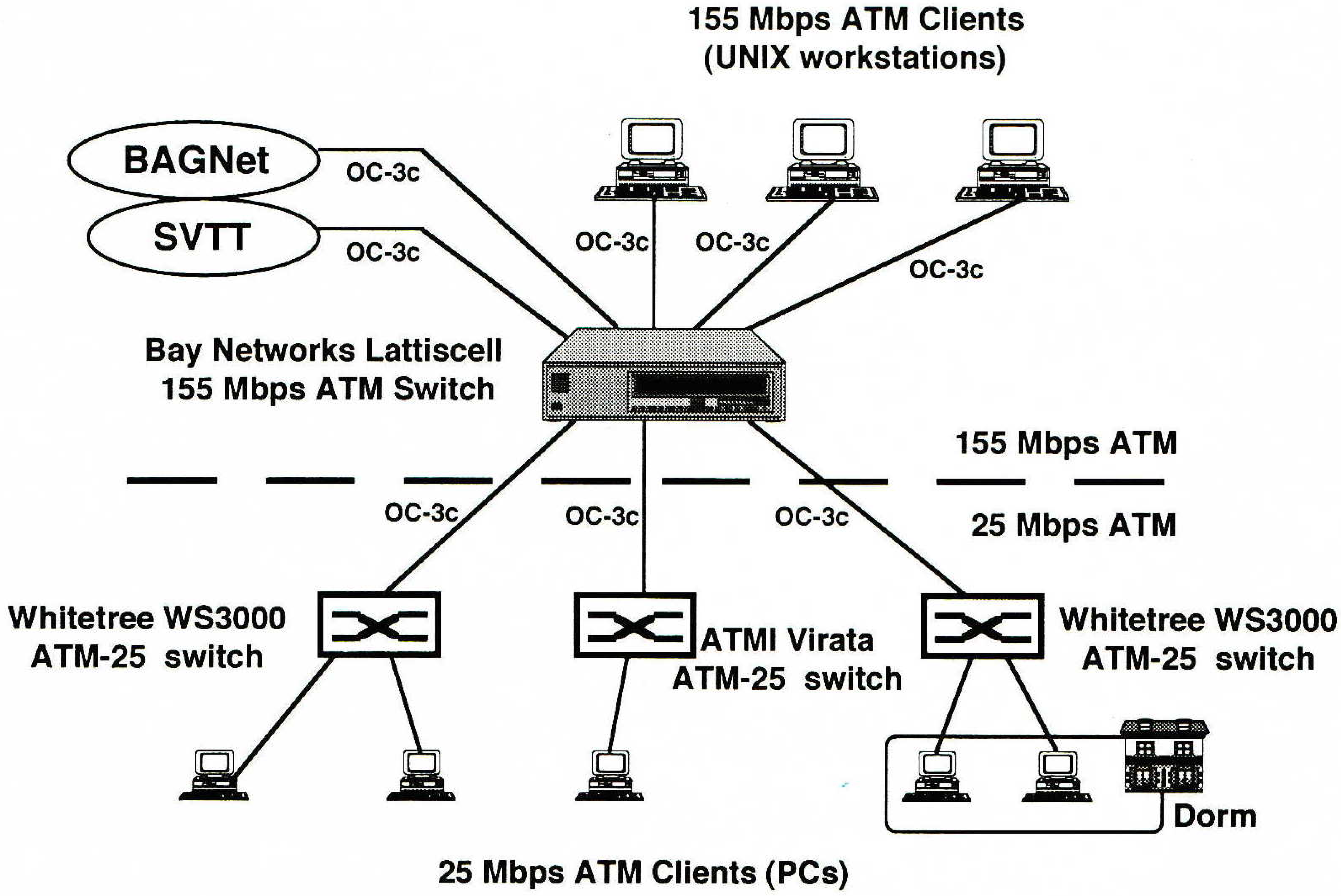

Figure 2: Stanford's experimental ATM network.

It is based on classical IP over ATM using logical link control/subnetwork attachment point (LLC/SNAP) encapsulation. It employs both User–Network Interface (UNI) 3.0/3.1 compliant Switched Virtual Circuits (SVCs) using Q.2931 signaling, and Permanent Virtual Circuits (PVCs) to interconnect end hosts. It also integrates a variety of ATM equipment from multiple vendors. At the heart of the network is a Bay Networks Lattiscell 10114R-SM ATM switch. A number of other ATM clients are also connected directly to the Lattiscell.

In addition to hosts with 155 Mb/s ATM access, the ATM network also has a number of PC hosts with 25 Mb/s ATM access. Three smaller ATM-25 switches are each connected to the Lattiscell over OC-3c connections for this purpose, as shown in the figure.

Many valuable technical lessons are being learned through ADEPT. The main lesson is that the delivery technology is finally coming of age, in terms of both capability and cost. Both multimedia technology and networking infrastructure is reaching critical mass. The Internet is still inadequate for delivering quality video in real time, but the emergence of broadband networks like BAGNet will solve this problem. Delivering distance education today, however, is still a complex undertaking requiring significant computing, networking, and skilled human resources to address multiple client platforms and heterogeneous networking environments. Cross-platform solutions provide the greatest flexibility and ease of implementation. Future ADEPT work will continue to examine real-time playback of video and audio over ATM networks. (Contact: Carlos Cordero, Stanford University, `cordero@isl.stanford.edu`).

***Project:* Distributed Scientific Collaboratories**

LBNL is working on demonstrations and experiments involving remote use of optical and electron microscopes over BAGNet (this is an example of "distributed scientific collaboratories"). (Contact: Bill Johnston, LBNL, `johnston@george.lbl.gov`).

***Project:* Featured on C|Net Central**

BAGNet was used by C|Net Central as a demonstration of what the MBone could be like in the future as bandwidth availability increases. C|Net aired this segment on the USA Network and Sci Fi cable channels over the weekend of August 5, 1995. A write-up of their segment is available at the C|Net Web server. (Contact: Lance Berc, Digital Equipment Corporation, `berc@src.dec.com`).

***Project:* High-Rate Data Transfer**

RIACS/NASA-Ames is developing a high-rate data-transfer protocol for transfer of large image files. Since many image-transfer applications can tolerate a low level of transmission errors, we are basing our protocol development on UDP rather than TCP. Our goal is to develop a data-transfer protocol that maximizes network throughput over ATM networks, while keeping transmission errors manageable. Of course, the level of transmission errors that is considered acceptable is application dependent.

We are currently experimenting with several techniques for data transfer, all of which attempt to keep the transmission pipe full. We are using multiple data streams, so that disk I/O, etc, can be overlapped with data transfer. Transmission errors are controlled via a low level of synchronization of sender/receiver activities.

Early results validate our approach. We are able to limit transmission errors to two to three percent, while achieving throughput rates that are several times higher than FTP rates. (Contact: Marjory Johnson, RIACS/NASA-Ames, `mjj@riacs.edu`).

continued on next page

BAGNet Experiences *(continued)*

***Project:* Interactive Terrain Navigation and High-Speed Network Storage**

LBNL and SRI have designed and implemented a wide-area network based, distributed-parallel data storage system ("DPSS") as part of an ARPA funded collaboration known as the MAGIC gigabit testbed, and as part of DoE's high speed distributed computing program. This technology has been successful in both the MAGIC and BAGNet environments, serving as a high speed front-end to various types of digital image libraries. The DPSS provides an economical, high performance, widely distributed, and highly scalable architecture for caching large amounts of data that can potentially be used by many different users and processes.

The "TerraVision" terrain visualization application allows a user to explore/navigate "real" landscape. TerraVision requests from the DPSS, in real-time, the sub-images needed to produce the view of the landscape. Typical use requires aggregated data streams of from 100 Mbits/s to 400 Mbits/sec that are supplied from several servers on the network.

TerraVision and the DPSS are routinely used in the BAGNet environment and demonstrate a prototype capability for interactive exploration of very large, distributed, on-line data sets. (Contact: Bill Johnston, LBNL, `johnston@george.lbl.gov`).

For more information see: `http://www-itg.lbl.gov/ISS` and `http://www.ai.sri.com/~magic/terravision.html`

***Project:* Performance Measurements**

Once BAGNet was working it was of interest to determine just how well this network performed. LLNL ran a number of performance tests utilizing *netperf* over BAGNet and within LLNL's ATM network. *Netperf* is a network performance tool that is capable of measuring the transactional (e.g., emulating NFS) and stream UDP/IP and TCP/IP throughput between two hosts. In all of these tests the measurements involved an off-the-shelf workstation running the TCP/IP protocol suite so these measurements must not be considered "ATM measurements." However, these performance measurements suggest how well (or poorly) current ATM products provide an integrated solution for a high-speed network of heterogeneous workstations.

There are many factors involved with network performance including the workstation architecture, operating system, TCP/IP implementation and configuration, host adapter software, and host adapter hardware. Although these performance tests often used BAGNet, they are testing the performance of the complete data path from the *netperf* client on one workstation to the *netperf* server on another workstation.

The details of all of these performance tests and graphs (approximately 50) of the results are available on the Web. Some observations derived from these performance tests will be presented here.

Cell pacing measurements

Tests were also performed using the HP ATM analyzer to measure how accurately ATM host adapter software and hardware perform cell pacing (i.e., bandwidth limited PVCs). Only two vendors were found that provided this capability. One host adapter provided the configured bandwidth up to about 10Mb/s, then delivered about 90% of the bandwidth at higher bandwidth configurations. The other host adapter delivered the configured bandwidth up to about 25Mb/s, then provided about 5% above the configured bandwidth.

Multicast performance

Three BAGNet workstations were selected for multicast tests. These hosts used *netperf* UDP/IP to perform tests using the same unicast IP addresses but two different PVCs; a unicast (point-to-point) PVC, and a multicast (point-to-multipoint) PVC with up to 14 branches. With the LLNL ATM switch, the unicast performance was indistinguishable from the multicast performance. Over BAGNet (3 ATM switches, one being the multicasting BAGNet backbone ATM switch) the performance was different: 1) The unicast PVC the maximum performance was 80Mb/s, while with the multicast PVC the maximum performance was about 20Mb/s, 2) With the multicast PVC, the throughput collapsed to zero with UDP datagrams greater than about 6KB. This did not occur with the unicast PVC.

UDP/IP transactional performance

The transactional performance (transactions per second), expected to be an indication of NFS performance, was quite good over ATM in general and also over BAGNet. This is not surprising since the latency measured over BAGNet via *ping* is quite low (~6 milliseconds). For example, the *netperf* transactional performance over BAGNet was 2–3 times that obtained between two LLNL hosts over FDDI and Ethernet when 2 or more routers were in the path, and comparable to two hosts on the same Ethernet segment (i.e., no routers or bridges).

TCP/IP performance

There were several interesting observations from the TCP/IP performance tests:

- *Inconsistent performance between hosts:* TCP/IP performance was observed to range from 20Mb/s to about theoretical maximum of 135Mb/s. For example, a Solaris 2.4 Sun SPARCstation 20 (SS20) topped at 70Mb/s to an IRIX 5.3 SGI Challenge, while the Challenge only achieved 38Mb/s to the SS20. Yet an SGI Indigo-2 and Solaris 2.4 SS5 both could transmit to the SS20 at about 60Mb/s.

- *Inconsistent performance with smaller buffers*: With socket buffers smaller than 20KB the performance low and often fluctuates wildly.

Surprisingly, the performance between Solaris 2.4 workstations was extremely consistent (approximately 45Mb/s) for all socket buffer sizes regardless of the ATM host adapter vendor (i.e., host adapters from FORE, Sun, and ZeitNet). It was observed by Alden Jackson of SNL that this behavior would occur if Solaris enforced a reasonably large (e.g., 16KB) minimum socket buffer size.

UDP/IP performance

The UDP/IP performance measured the throughput of the receiver. There is nothing throttling the sender with UDP as there is with TCP, so fast senders would overwhelm slow receivers. Some observations of the UDP/IP performance:

- Some receivers measured a throughput of practically zero from certain senders, while some receivers did quite well, achieving over 100Mb/s.

- Slower senders generally resulted in consistent (but mediocre) throughput for most receivers, about 50Mb/s.

Summary

Considering its maturity level, ATM's performance appears to be quite good. For example, with a Solaris 2.3 SPARCstation 5 (the only workstation available to me with all these media), ATM compares favorably with other *netperf* TCP/IP measurements obtained at LLNL (in round numbers): 40Mb/s for ATM, 40Mb/s for Fibre Channel and 25Mb/s for HIPPI, which are 155Mb/s, 266Mb/s and 800Mb/s media respectively.

continued on next page

BAGNet Experiences *(continued)*

Newer ATM host adapters and software are providing much better performance, with reports of several workstations now capable of achieving the maximum TCP/IP throughput over ATM (~135Mb/s).

These and other experiences with ATM over wide-area networks emphasize the importance of larger socket buffers for applications that use TCP/IP (e.g., FTP). Workstation and PC vendors should be addressing such application level network issues, particularly now that high-speed, wide-area networks are becoming more evident. (Contact: Dave Wiltzius, LLNL, `wiltzius@llnl.gov`).

***Project:* Public ATM Networks for Health Care Imaging Systems**

We (W. Johnston, LBNL; E-J. Pol, Philips Research; and J. Terdiman, MD, Kaiser Permanente Division of Research) are exploring the use of a shared ATM network to facilitate collection, storage, analysis, distribution, and delivery of several kinds of health care imaging data. The data includes still images and video sequences of coronary angiograms. The data is being collected by a computer system in a cardiac catheterization laboratory, and that system is attached to a Pacific*Bell, OC-3, metropolitan area ATM network. The image data is stored on network-based multimedia storage systems and subsequently delivered to physicians at other locations around the SF Bay Area.

The project is characterizing and addressing various issues, including the ability of the ATM network to deliver the data end-to-end in a useful way, the security of the data, the availability of network resources, the ability of state-of-the-art computing systems and software that are interconnected by ATM networks to perform at the levels required for routine clinical use, etc. Apart from the basic issues of data collection, transport, and delivery, the project will explore the use of the network to provide flexible data storage strategies and location independent access to analysis systems. (Contact: W. Johnston, LBNL, `johnston@george.lbl.gov`).

***Project:* Remote Access to NASA Wind Tunnel**

NASA Ames is experimenting with the use of networks to enable remote control of wind-tunnel experiments. This is part of a major, on-going project at Ames to provide remote access for aircraft designers to the wind-tunnel control room, so that they can monitor and control their experiments without having to be physically present while they are running. BAGNet is being used to evaluate the use of ATM networks for this application, as compared to the use of other network technologies. (Contact: Alfred Nothaft, NASA Ames, `nothaft@nas.nasa.gov`).

***Project:* Supercomputing 95**

LLNL had a unique opportunity to extend a few BAGNet hosts to the Supercomputing 95 (SC95) conference in San Diego in December 1995. ESnet (Energy Sciences network) obtained an OC-3c ATM connection from LLNL to the SC95 conference floor from Sprint. Working with Sprint and the ESnet network team, LLNL obtained a virtual path between LLNL and the LLNL booth at SC95. Several of LLNL's BAGNet PVCs were then reconfigured through this virtual path and to several ATM hosts at the booth. Stanford and LLNL demonstrated distance learning capabilities (such as MPEG-1 audio/ video to workstations and PCs) and applications over this network at SC95. Although this was one IP "hop" these ATM virtual circuits traversed an estimated 13 ATM switches, 6 at SC95 and the rest at LLNL and on BAGNet.

LBNL used this infrastructure at SC95 to remotely control and view results from a special optical microscope and a micro-spectrometer at the Advanced Light Source at LBNL.

During this conference we noted problems with the quality of the MPEG-1 received by the Pentium PCs that were not noticed on BAGNet. Follow-up investigations at LLNL with Stanford and Sandia (and their long-link emulator) demonstrated that the problem was latency, due to the longer distance of the ATM network at SC95, which throttled TCP/IP. This phenomenon is well understood (e.g., bandwidth-delay product [10]). We suspect the PC's default socket buffer size is unusually small (approximately 2.5KB vs 8–64KB for workstations). (Contact: Dave Wiltzius, LLNL, `wiltzius@llnl.gov`).

***Project:* TCP/IP Throughput Performance Variance**

Given the TCP/IP performance measurements and interesting behavior found by Dave Wiltzius at LLNL, additional TCP/IP throughput experiments with cell pacing were performed to examine the distribution of the experimental measurements. These studies would show if the performance between two nodes is generally poor at certain parameters, or if the performance exhibits a wide range of values. If the performance for a fixed set of parameters is generally poor for one set of hosts, but exchanging one of the hosts gave better performance, then the culprit is probably in the host. If the performance for a fixed set of parameters varies widely from measurement to measurement, it suggests an instability in the system that could come from several sources, including the interaction of several sources. Additionally, the experiments included BAGNet, allowing the expansion of the number of host to host pairs tested.

Netperf (v2.0) was the tool used to test TCP/IP throughput performance between BAGNet hosts. The measurements used standard workstations and vendor-supplied TCP/IP implementations over BAGNet.

In general, multiple experiments were performed on a range of socket buffer sizes (8192 to 131072 bytes) and a range of cell pacing bandwidths (70 Mb/s to available bit rate (ABR) at OC-3). For each parameter set, identified by socket buffer size and cell pacing bandwidth, the multiple measurements were sorted, the median and mean extracted as well as the minimum, maximum, and lower and upper quartiles. If the mean and median are close in value then the experimental measurements are equally distributed about the mean. If the mean and median are not close in value, then the measurements are not equally distributed about the mean. This result can occur if the measurements are grouped about certain values, or the distribution is not symmetric. The details of the experiment design and results are available on the Web at (`http://auditorium-ether.ca.sandia.gov/`). What follows are some observations from those experiments.

Observations

- These tests corroborated the inconsistent performance Wiltzius found with buffers less than 50KB. In this region, the performance is characterized by severe non-linearity. There is also very little variance in the measurements for any of the cell pacing bandwidths in this region.
- For buffer sizes greater than 50KB and cell pacing bandwidth less than 100 Mb/s there was very little variance in the measurements. It was noticed that the performance was dependent on the host pairs. For example, between DEC Alpha 3000/600's the throughput measured consistently approached the cell pacing limits.

continued on next page

BAGNet Experiences *(continued)*

Yet between a DEC Alpha and a Sun SS20, the maximum throughput was about 75 Mb/s, and throughput measured at cell pacing limits lower than 75 Mb/s were effected also.

- For buffer sizes greater than 50KB and cell pacing bandwidth greater than and equal to 100 Mb/s, there was a large variance in the measurements between hosts that can support high throughput. For example, the ABR performance when not limited by buffer size between DEC Alpha 3000/600's ranged from 133 Mb/s to less than 50 Mb/s. Although 75% of the measurements were above 100 Mb/s, 25% of the measurements were below 100 Mb/s and 10% below 60 Mb/s. Host pairs that are bandwidth limited did not exhibit this behavior.
- Cell pacing does have its benefits. Although there is variance in the measured throughput for buffer sizes greater than 50KB and cell pacing bandwidth greater than and equal to 100 Mb/s, the variance decreases as the cell pacing bandwidth becomes smaller.

SNL is continuing to investigate why this behavior occurs. The SNL performance Web page mentioned above (and in the URLs at the end of this article) will document future research. (Contact: Alden W. Jackson, SNL, `awjacks@ca.sandia.gov`).

***Project:* Video-on-Demand**

NASA-Ames has several VOD efforts, serving up MPEG-1 and MPEG-2 over wide-area ATM networks. (Contact: David Meyers, NASA-Ames, `dmeyers@vod.arc.nasa.gov`).

***Project:* BayNet**

Recently we had Pac*Bell configure several virtual paths between some sites on BayNet (another CalREN funded Pac*Bell ATM testbed, in the Monterey Bay area, see page 22). We have a 15Mb/s virtual path between the University of California Santa Cruz (UCSC) campus and the extension in Santa Clara and are running both FORE SPANS and UNI 3.0 signaling protocols to do SVCs over this virtual path. This made it very easy to add machines on our end (e.g., two SGIs were added to our local switch and shortly thereafter we had full connectivity between the sites).

We transmitted a class over BayNet this last quarter, using SPARCstation 20s equipped with Parallax video cards for transmission and reception.

Another virtual path was configured between UCSC and the Naval Postgraduate School (NPS) in Monterey. Our plan is to add sites one at a time, and deal with the bandwidth allocation issues as they arise, rather than trying to map the entire topology beforehand (BayNet ATM resources are configured as a "guaranteed service," where the VPCs and VPs are bandwidth limited—unlike BAGNet). We will be using the UNI 3.0 signaling only (i.e., no SPANS) as the NPS has Cisco equipment. Our first pass will probably entail having both UCSC and NPS on the same subnet with our switch acting as the ATM ARP server. We will break things up into subnets once the network is working, so the NPS doesn't have to depend on our switch for setting up their connections.

UCSC also showed a collaborative visualization demo at SC95 between two SGIs, one at UCSC and the other at the SC95 conference floor in San Diego.

We are also looking into other ATM codec equipment to see if we can provide full screen video for the lectures. (Contact: Arul Ananthanarayanan, University of California, Santa Cruz, `arul@cse.ucsc.edu`).

Conclusion

BAGNet was an interesting learning experience for many of us involved with this project. We hope that the information we presented will help others better understand issues pertaining to ATM metropolitan-area networks at this early stage of deployment.

Acknowledgements

We would like to thank Pacific*Bell for providing infrastructure and support for this network testbed consisting of such distinguished organizations. Thanks also to Rick Hronicek for his work to establish CalREN in such a way that BAGNet could be funded, and to Bob Kahn who provided the initial impetus. Special thanks to Pacific* Bell's BAGNet Project Leader Hon So for his efforts throughout the project, and the technicians in Pacific*Bell's Network Operation Center for their upbeat, professional and tireless assistance. We would also like to thank the several BAGNet participants that helped review this article and contributed material.

URLs of interest

BAGNet home page:
`http://www-itg.lbl.gov/BAGNet.html`

BayNet home page:
`http://www.cse.ucsc.edu/research/baynet-atm/`

BAGNet ATM performance:
`http://www.llnl.gov/bagnet/perf-llnl.html`

BAGNet problem analyzed (35 GIF images):
`http://chocolate.research.digital.com/hoti95/talk.html#mystery`

BAGNet on C|Net:
`http://www.cnet.com/Content/Features/Techno/Mbone/index.html`

BAGNet at CompCon:
`http://www-itg.lbl.gov/BAGNet.hm.pg.docs/CompCon.95.3.html`

MBone information:
`http://www.eit.com/techinfo/mbone/mbone.html`

References

[1] M. Laubach, "IP Over ATM and the Construction of High-Speed Subnet Backbones," *ConneXions,* Volume 8, No. 7, July 1994.

[2] T. Wu, *Fiber Network Service Survivability,* Artech House, 1992.

[3] B. Amin-Salehi, G. Flinchbaugh and L. Pate, "Implications of New Network Services on BISDN Capabilities," Proceedings of IEEE Infocom '90, pp. 1038–1045, June 1990.

[4] "Some Considerations in Selecting a Cell Information Field Size," Bellcore, Tech. Report T1D1, 1/87-493, 1987.

[5] S. Minzer, "Broadband ISDN and Asynchronous Transfer Mode (ATM)," *IEEE Communications,* Volume 27, September 1989.

[6] M. de Prycker, *Asynchronous Transfer Mode—Solution for Broadband ISDN,* 2nd Ed, Alcatel Bell, 1993.

[7] R. Händel and M. Huber, *Integrated Broadband Networks,* Addison-Wesley, 1991.

[8] D. McDysan and D. Spohn, *ATM: Theory and Application,* McGraw Hill, 1995.

[9] "ATM Services, The Truth Behind the Hype," *Telephony,* September 1995.

[10] W. R. Stevens, *TCP/IP Illustrated, Volume 1: The Protocols,* Addison-Wesley, 1994.

DAVE P. WILTZIUS is a member of Computations staff at Lawrence Livermore National Lab (LLNL). He has been frequently involved with high-performance network issues and applications while at LLNL. He has a B.S. in Applied Math and Physics and a M.S. in Computer Science from the University of Wisconsin-Madison. E-mail: `wiltzius@llnl.gov`

LANCE BERC works for Digital Equipment Corporation at the Palo Alto Systems Research Center. On weekends with favorable conditions he can often be found skiing or sailing, leaving his Alpha available for BAGNet performance testing. E-mail: `berc@src.dec.com`

SIDDHARTHA Y. DEVADHAR (Sidd) is with the Network Services Group in Pacific Bell. He has a B.Tech. (Mech) from IIT, Bombay, 1985, M. E. (Mech) Tulane Univ., 1987, and M.S.(CS) 1993. His interests are traffic modeling, synthetic traffic generation and trace-driven simulation and performance analysis of switches. E-mail: `sydevad@srv.pacbell.com`

Monterey BayNet

by
Craig M. Wittenbrink,
Arul Ananthanarayanan,
J.J. Garcia-Luna-Aceves,
Patrick Mantey,
University of California, Santa Cruz
and
Bruce Gritton, Monterey Bay Aquarium Research Institute

Introduction

The *Monterey BayNet ATM Project* is one of the CalREN funded testbeds. The proposal was to use ATM as an enabling technology for tele-education, tele-science, and electronic libraries linking the Silicon Valley, Santa Cruz, and The Monterey Peninsula. J. J. Garcia-Luna and Patrick Mantey (co-authors of this article) wrote the proposal with the backing of many regional collaborators including the UCSC Chancellor Karl S. Pister, The Tech Museum of Innovation's president, Peter B. Giles, the Superintendent of the Naval Postgraduate School, T. A. Mercer, the executive director of the Monterey Bay Aquarium Research Institute, Peter G. Brewer, and the executive director of the Monterey Bay Aquarium, Julie Packard.

The challenge was to weave a multidisciplinary regional infrastructure using the strengths of the participants. The CalREN funded ATM network is just a single component in a wider push made by regional collaborators which has resulted in funding for a frame relay network, virtual field trips educational development and evaluation, and partnered funding of continued network and content development such as that committed by the Monterey Bay Aquarium—due to the success of BayNet. Ever aggressive goals included:

- A new educational paradigm (interactive, exploratory, current, distance independent), and life-long learning opportunities for the 21st century schools, government, and industry of the Silicon Valley and Monterey Bay region.
- Ubiquitous access and timely delivery of environmental and oceanographic information to users in the various economic sectors of the Silicon Valley and the Monterey Bay region.
- Innovative information products and services that forge new linkages and collaborations between economic sectors and geographic regions.
- Dynamic dissemination mechanisms for providers of public information products and services.

This article discusses the status to date of BayNet: applications, infrastructure, and participation, and the near term plans with the remaining one year of the grant. Transition plans to other support for regional networking are also discussed.

Monterey BayNet-ATM

The Monterey BayNet-ATM network consists of seven sites in two telephone LATAs. We use "BayNet-ATM" to distinguish this effort from the collaborative regional effort of BayNet-Educational Frame-Relay network. The locations of the nodes of the ATM network are:

- The University of California, Santa Cruz (UCSC)
- The University of California Extension Santa Cruz (in Santa Clara)
- The Tech Museum of Innovation in San Jose
- The Monterey Bay Aquarium Research Institute (MBARI)

- The Monterey bay Aquarium (MBA)
- California State University Monterey Bay (CSUMB-Fort Ord)
- The Naval Postgraduate School (NPS)

Figure 1 shows the physical topology of the network. Note that an intermediate Sprint switch is required because the network spans the LATA boundaries legislated by the state. The Pac*Bell switches in both LATAs are shared with other applications, for example the Palo Alto switch is also used in the fabric of BAGNet [1]. Such interconnections have held promise for interconnecting the networks, and it was planned to have the University of California Berkeley, as part of BayNet.

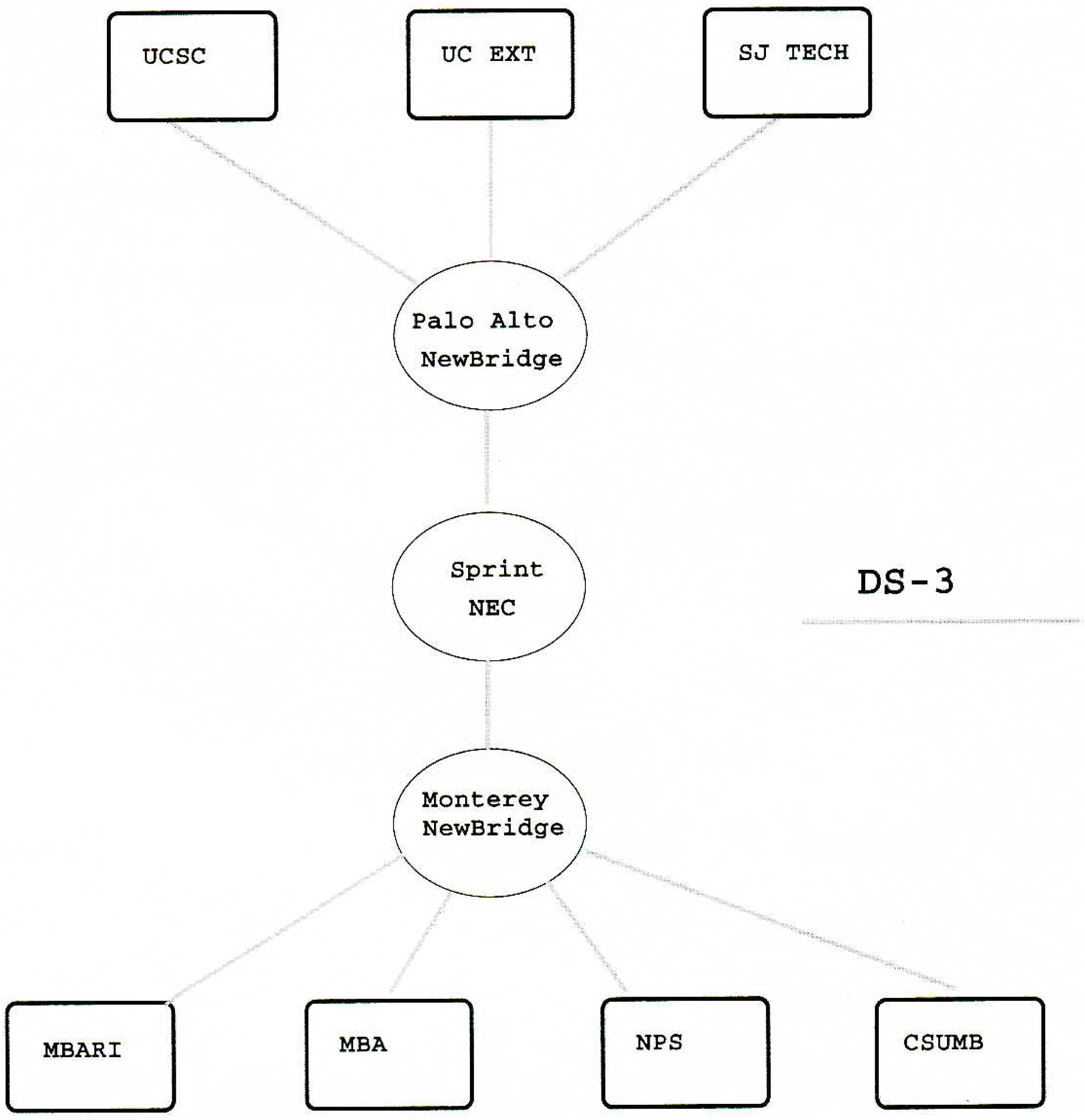

Figure 1: Physical BayNet ATM Switch Fabric

While Pacific*Bell donated nearly $25 million to operate the California Research and Education Networks (CalREN), these funds allocated only the bandwidth and Pacific*Bell's necessary infrastructure. For our seven sites, we needed to come up with independent funds for switches, workstations, pulling fiber, ATM workstation adapters, and video/audio equipment. Important to making this happen in a cost effective manner for our network of educational and nonprofit institutions were the alliances with commercial participants for group buys. Commercial participants include: Ask/Ingres (now CA), Cisco, FORE, HP, Insoft, Newbridge, Newman and Associates, Paradise, Parallax, Sun Microsystems, and STS Corporation.

The BayNet network became operational in 1994 with the initial sites of UCSC, MBARI, and MBA. Many applications were planned, but the highest visibility application is BayLink. BayLink provides the content from MBARI's remotely operated vehicle to the San Jose Museum of Technology. Other applications, including UCSC to UCSC Extension distance learning and CSpray—a collaborative scientific visualization application—have been brought up in the last year.

continued on next page

Monterey BayNet *(continued)*

Because the network is the collective result of many disparate organizations, and funding availabilities, the technological aspects of the network have proved to be a minor part of creating BayNet. We will go into further details of the ATM research, demonstrations, and applications in BayNet, as well as discuss the issues in a regional collaboration.

Design

In a companion article [1], Dave Wiltzius et al. give an introduction to ATM. ATM networks involve a variety of layers. ATM is a collection of standards for communications in broadband optical networks with integrated services of voice, video, and data traffic. In Monterey BayNet-ATM, the backbone network consists of DS-3 links, providing 45 Mb/s. Within each site links used are OC-3, and provide 155 Mb/s. The primary difference between Monterey BayNet-ATM and BAGNet is the provisioning of guaranteed service instead of best effort service.

The physical topology could be the logical topology if switched virtual circuits (SVCs) were available at the backbone. Similar to BAGNet, the backbone switches did not provide SVC capability. Our initial topology consisted of permanent virtual circuits (PVCs) as shown in Figure 2. Due to the interLATA boundary, and the limited bandwidth, PVCs were carefully chosen to accommodate the video applications for BayLink and the Distance learning.

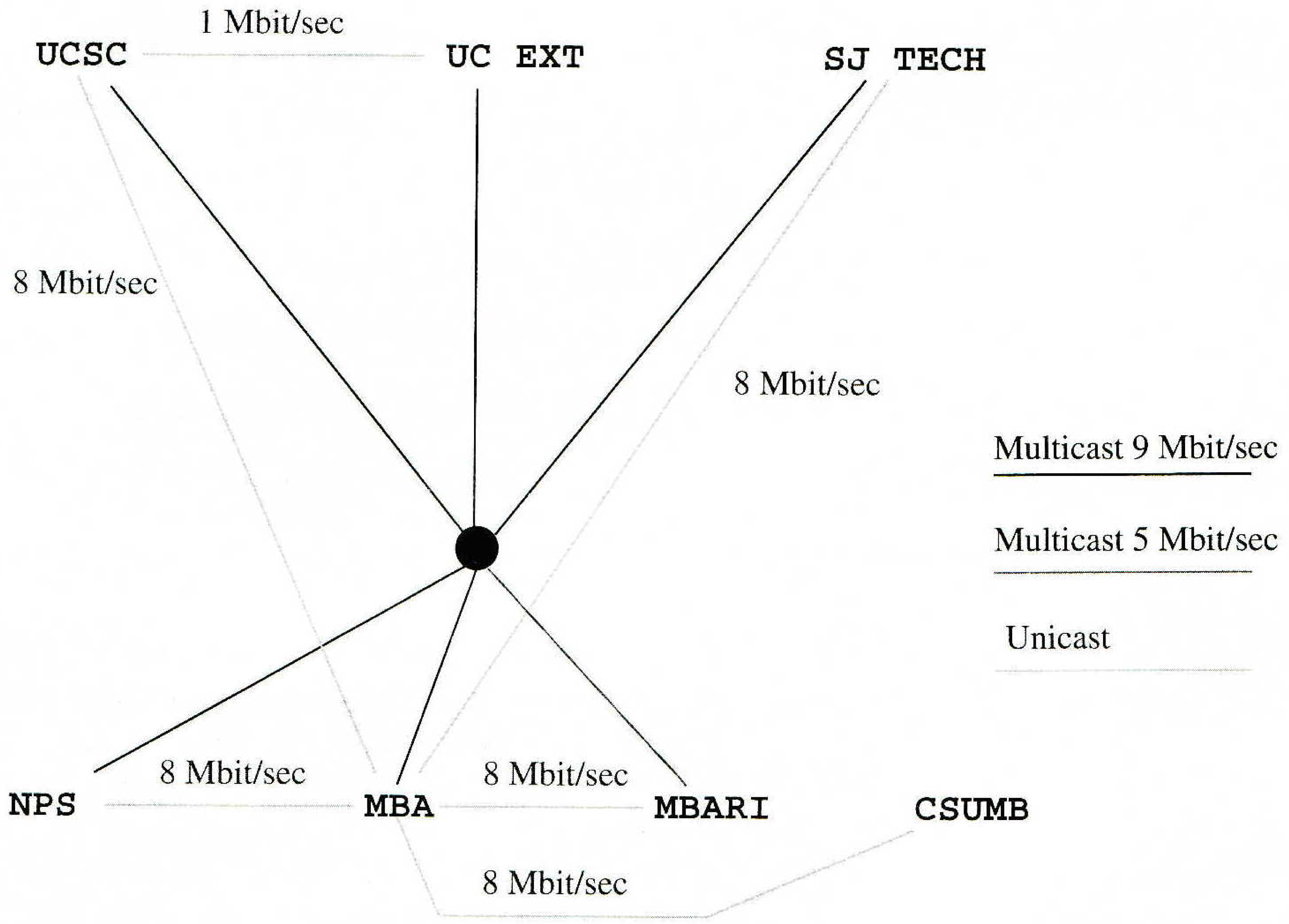

Figure 2. BayNet ATM Initial PVC Logical Topology

However, as the project progressed, we decided to migrate to a network infrastructure based on virtual paths, for the following reasons:

- The rate allocation assumed that everyone was using their entire allocated bandwidth all the time.
- Only a couple of the BayNet-ATM sites were on line at the same time, so most of the bandwidth was not being used.
- The rate allocation provided us with a 1 Mb/s PVC to the UC Extension and this was clearly too small to accommodate our distance learning application.

- The PVC backbone and the Point to Multipoint mesh that was originally allocated proved to be difficult to manage because of the number of PVCs and the fact that network sites had to contact Pac*Bell every time changes had to be made to the topology. Adding new hosts to the sites would make the PVC mesh even more complex, or would introduce complicated local routing methods.
- Our old network topology used an approach similar to the other CalREN projects (e.g., BAGNet) [1].

The BayNet ATM project was the first CalREN project to use virtual paths. Once the UC Extension site became operational, we obtained a 15 Mb/s Virtual Path between the UC extension and UC Santa Cruz. This path provided enough bandwidth to run our distance learning application between UCSC and the UC Extension, and still allowed other BayNet ATM applications to run, BayLink in particular.

Currently, we are running FORE's proprietary SPANS protocol over the virtual path. This allows us to use SVCs to communicate between machines here and at the Extension. All of the SVC setup is handled by the switches at either end. FORE's SPANS protocol also provides native multicast support, so we can do multicasting without setting up a complicated point-to-multicast PVC mesh. The switches will do all the work for us.

We are also running a Classical IP network over the virtual path. This allows us to use SVCs using the Q.2931 signaling protocol described in the ATM Forum UNI 3.0 specification. Classical IP over ATM is described in RFC 1577 and is supported by all the major ATM switch and adaptor vendors. This means that sites using FORE equipment can interconnect with other BayNet sites that are not using FORE equipment.

The signaling protocols were set up very quickly and smoothly over the virtual path between UCSC and the UC Extension site. This was possible because of all the testing work done locally over the summer at UCSC.

The virtual path allows our project to manage its own bandwidth and VCs. BayNet ATM can still setup PVCs within the virtual path, but user sites do not need to contact Pac*Bell to make these changes.

Our current network setup consists of four machines at UCSC (SPARC 5, SPARC 20, SGI Indy, and SGI Indigo^2 Extreme) hooked up to a FORE ASX-200 switch. We have a FORE ASX-200 switch and a SPARC 20 at the UC Extension. The virtual path scheme has been extended to NPS, where testing is in progress. The manual setup of the virtual path from UCSC to NPS has been requested, but has not gone through Pacific*Bell yet.

Applications and requirements

BayNet is being used for a number of telescience, tele-education, and electronic library applications being developed at the participating sites. The applications include:

- BayLink video from MBARI to MBA and The Tech Museum.
- UCSC to UCSC-Extension distance learning.
- Collaborative Scientific visualization of weather and oceanographic data with UCSC's CSpray software.
- REINAS—*Realtime Environmental Information Network and Analysis System*—for environmental data archive and access, as well as remote site video services.

continued on next page

Monterey BayNet *(continued)*

BayLink application

The Monterey Bay region has the opportunity to get our message to a larger world. A message dominated by the theme of people living and working in harmony with our environment through publicly visible exploration, research, and education of the Monterey Bay Submarine Canyon ecosystem. By taking advantage of ATM, video compression hardware, and distributed multimedia software, we have extended the MBA/MBARI Live Link to The Technology Museum of Innovation in San Jose.

This capability allows students at The Tech to participate in the exploration of the Monterey Canyon by MBARI scientists as it happens. Real time, full motion video from the bottom of the canyon is supplemented by expert interpretation from the MBA. Just as the audience in the MBA auditorium can interact with the interpreter (called the LINKER), a participant in San Jose views and addresses the LINKER and directs questions to the scientists in the field. Deployment of BayLink is planned for extended distribution to other classrooms/museums around the world.

A quality requirement for BayLink was true 30 frames/second with 640 × 480 pixel resolution. The MBA was very wary of the technology, and any reduced frame rate or reduced resolution solutions, were unacceptable. Craig Wittenbrink (one of the authors) and Mike Newman evaluated a large number of available compression hardware/software, and selected the Parallax SVIO motion JPEG compression boards for Sun SPARC SBus. Other JPEG cards included DEC Alpha, Parallax for HP, HP video, and SGI Galileo, but none offered demonstrated 30 fps networked color video of the required resolution. Since that time dedicated ATM audio/video codecs, such as the unit from STS Corporation, and the one from Nemesys Research have become available. A difficulty with the Parallax hardware has been drivers, and unfortunately we had to resort to using two workstations at each BayLink Site, as only one Parallax card could work in one workstation. As the digital video market is maturing, current (1996) solutions would be more cost effective using native ATM audio/video codecs, or going with standards based H.320 solutions.

UCSC to UCSC-Extension Distance Learning

For this application, we have studied many aspects of the distance-education process, with the intent of making distance learning an integral part of the way in which UCSC and the UC Extension service their community. With the knowledge gained from evaluating equipment for BayLink, the decision was made to go with the Sun/Parallax hardware. For the distance learning, we evaluated two collaborative software products: Insoft's Communique and Paradise's PSVC—both running Sun SPARC 20 on Parallax. For our distance learning application we also use remotely controlled auto focusing cameras (Canon VC-C1) and high resolution 1280 × 1040 Sun to NTSC scan converters for local large video monitors (PCVideo Corp). The video solution in this case was to allow half the frame rate (15 fps) therefore requiring only a single SPARC 20 at each site, and 1 Parallax card at each site. We have found PSVC to be much simpler to set up, but Communique to have more features. Most of the classes were distributed on PSVC. The major difficulty for any collaborative software is audio feedback (echo cancellation) and floor control. Classroom issues such as whiteboards, microphone protocols, and timely technical support were more important than the ATM and digital video/audio solutions. UCSC is actively working on developing distance learning classes and support.

The following equipment is being used at UCSC and the Extension. This equipment is very similar to that used on BayLink:

- Sun SPARC 20 running Solaris 2.4
- FORE SBA200 ATM interface.
- Parallax Video card
- PC Video Conversion Inc. scan converter
- Canon video camera
- Microphone
- Television
- Insoft Communique
- Paradise Simplicity software (PSVC)

The Suns communicate with each other over the virtual path using an SVC. The Paradise (PSVC) or Insoft software (Communique) use hardware motion-JPEG of the Parallax, XVideo card for capture, compression, and decompression. The video conferencing software uses the built in audio chip on the SPARC 20 to sample and decompress audio.

The scan converter allows full screen projection of the the workstation monitor to support digital presentations, electronic white-board, slides, captured image display, and demonstration of graphics and visualization software over the ATM link.

CSpray Collaborative Visualization

We have developed middleware in CSpray. CSpray is a computer supported cooperative work (CSCW) application geared towards supporting multiple users in a collaborative scientific visualization setting. Scientists are allowed to share data sets, graphics primitives, images, and create visualization products within a view independent shared workspace. CSpray supports incremental updates to reduce network traffic, separates large data streams from smaller command streams with a two level communication strategy, enforces permissions for different levels of sharing, distinguishes private from public resources; and provides multiple fair and intuitive floor control schemes for shared objects. Off-the-shelf multimedia tools such as *nv* and *vat* can be used concurrently. CSpray is based on the *spray* rendering visualization interaction techniques [9] to generate contours, surfaces, particles, and other graphics primitives from scientific data sets such as those found in oceanography and meteorology.

We use supercomputer calculated forecast models from the Naval Postgraduate School provided by Wendell Nuss, collaborator in the Department of Meteorology. The tool aids meteorological professionals in explaining, understanding, and sharing up-to-the-minute weather and storm analyses. ATM provides the multiple services necessary to fuse the video conferencing information with scientific visualizations. CSpray provides a multiparty cooperative visualization session, where any user can contribute or probe remote data. ATM creates the ability to analyze larger datasets interactively, and to match the rendering speeds of high end Silicon Graphics workstations for the most effective displays. Data are not transferred, as the ATM connected computers serve as a heterogeneous database servicing requests from each site, and providing the highest quality visualization products. The visualization products, namely graphics primitives can be exchanged and rendered locally taking maximum benefit from the DS-3 data transfer rates available.

continued on next page

Monterey BayNet *(continued)*

BayLink demonstration

On May 5, 1995 the BayLink application was demonstrated at the Department of Commerce auditorium in Washington D.C. as part of the National Information Infrastructure Task Force meeting. Two high schools, one in North Carolina and one in Illinois, participated in real time as science was conducted at the bottom of the Monterey Bay Canyon. We are very pleased that the BayLink application was one of the first demonstrations of a transcontinental, end-to-end ATM application, and that it was recognized by the President of the U.S.

Since that time the application has been demonstrated at Supercomputing 95. The application has moved to a new suite of hardware workstations and appropriate changes have been made to take advantage of bug fixes available in the video hardware, software, and the Pac*Bell switch fabric. BayLink between the Monterey Bay Aquarium and The Tech Museum is operational. There have been training and acceptance issues at both the aquarium and the Tech, but the success of BayLink has resulted in the Monterey Aquarium pledging new resources to distance learning. Other distance learning applications for the Monterey Bay Aquarium and MBARI are planned, and the demonstrations have fueled development of new applications of the ATM network.

CSpray demonstration

Collaborative visualization using CSpray over the BayNet ATM network was demonstrated at Supercomputing 95. Two Indigo^2 Extremes running CSpray were connected through Pacific*Bell's network from The convention center in San Diego, to the UCSC's graphics laboratory. Continuous demonstrations were given for four days to attendees of Supercomputing 95 from within the Pacific*Bell ATM/Frame Relay demonstration booth.

The first experiment of distance learning between UCSC and the UC Extension site was based on the graduate course Econ 200A. The class ran on Mondays and Wednesdays from 4:00pm to 5:45pm. We are using between 7 and 8 Mb/s for this application, and we obtain between 12 and 14 frames/sec for video. Additional courses are planned, but due to the limited room size classes are not planned until Spring 1996 (when a new room is available).

Collaboration

The large number of collaborating partners in BayNet has been very productive, but also a difficulty for many of the organizations. As ATM and broadband technologies provide new capabilities, organizations must come up with new applications for them. This requires personnel and equipment commitments from all of the participating organizations. Much of the lag in developing applications and installing infrastructure has been due to the lack of commitment. Probably the most fruitful piece of the collaboration is—those people who become involved do so for their own interests and actively fight for resources and results. The buy-in by participants to a shared vision or goal cannot be underestimated in a multigroup endeavor. The partnered success of other regional related efforts, especially the Educational network directed by Kam Matray with support from network design team many of whom were NPS students illustrates precisely the University/Primary education collaboration desired. MBA has also recouped a great amount of its effort spearheaded by Bruce Gritton (one of the authors). The successful Washington D.C. demonstration, Tech Museum display, and several conference demos has given MBA an extended reach. MBARI's chairman of the board (Packard) has challenged MBARI/MBA to provide its audio/video content as widely as possible through broadband technologies such as the BayNet ATM network.

The regional networking infrastructure and collaboration is looking to the future, with the Futures Network director Bruce Gritton (one of the authors). Planned transition of BayNet beyond the CalREN funding and wider reaching networks are in the agenda for the Futures Network. And of course, the participation of many users has allowed UCSC to develop applications that are forward looking and useful in broadband network collaboration.

Research and future directions

BayNet is being used for a number of research projects as well as our applications to date. Our research and development work in this project includes:

- Installation and testing of ATM switches and hardware at the participating sites. All but one of the ATM sites are on-line.
- Development of lecture material for tele-education demonstration.
- Development of floor control and multicasting support for multimedia conferencing. This work is being carried out at UCSC as part of the REINAS project.
- Research on congestion control in ATM networks.

BARBARA

We hope to be able to partner with Pacific*Bell to use ATM service to provide the high-bandwidth necessary for the University of California, Santa Cruz Monterey *Bay Area Regional Broadband Archive and Real-time Access* (BARBARA). Multidisciplinary researchers and students may benefit from real-time weather data access and on-line video, audio, and visualization weather briefings. At UCSC, this project is complemented by the research carried out in the REINAS project, the Computer Communication Research Group (CCRG), the High-Speed Networks Lab, and the Santa Cruz Laboratory for Visualization and Graphics (SLVG) of the Baskin Center for Computer Engineering and Computer and Information Sciences. This research is being funded primarily by grants from ARPA/ITO, ONR. For details on the system components, demonstrations, and research being done in REINAS, you can refer to the REINAS Home Page.

NPS future research

At NPS, work is underway by graduate students at the NPS Information Infrastructure Research Group (IIRG). A study seminar (CS 4920-2) on ATM was run to discuss the issues involved in developing ATM networks and applications.

Future network designs

Given the success in our Virtual Path experiments, we are planning to connect UCSC to NPS in the same way by extending another virtual path to the NPS site. We will either designate the ATM switch at UCSC as the ARP server, so that NPS can join our Classical IP network, or we will have to use another subnet to allow NPS to use its own ATM switch as a server. In the process, we will apply for a Class C address and set up routing on the ATM switches. Our plan is to extend the BayNet ATM network using VPs on a site by site basis and deal with the complexity as we go along. We are hoping that we can allocate the entire DS-3 bandwidth for the virtual path and use rate control on the individual adaptors or switches to make sure that the available bandwidth is not exceeded.

As part of our virtual path experiment, we're planning to create a ring topology (see Figure 3) to provide full connectivity between all of the BayNet sites. This is due primarily to the limitation of no SVCs in the backbone switches (else the tree would be more optimal). The ring topology will allow each site to allocate their entire DS-3 bandwidth.

continued on next page

Monterey BayNet *(continued)*

Our original intention was to create a virtual path mesh between all of the sites, but we decided against this because the resultant bandwidth available for communicating between individual sites would not have been sufficient for our application requirements.

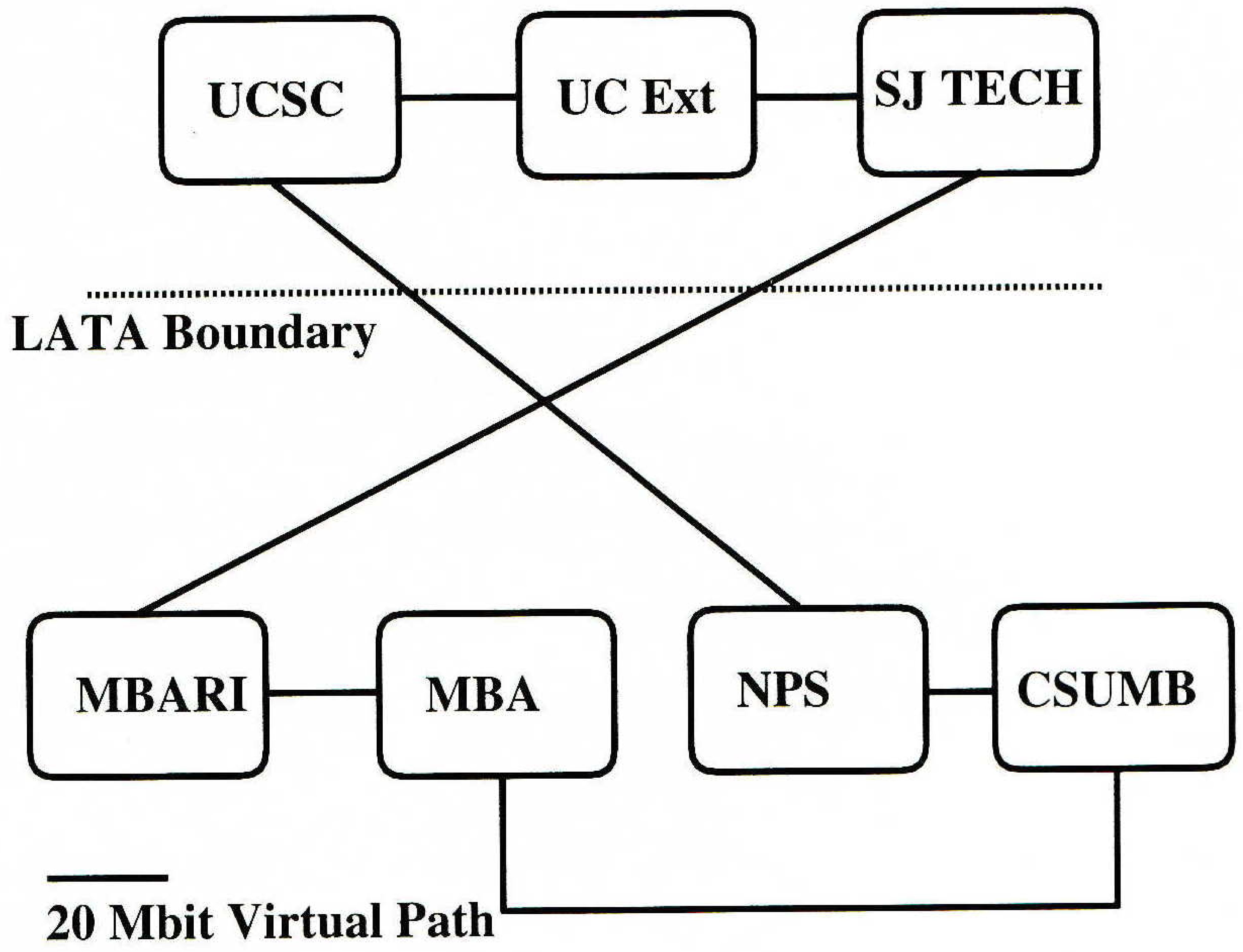

Figure 3. BayNet ATM Virtual Path Logical Topology

An important part of the rest of our development on the ATM network infrastructure is accomplishing multicasting among the BayNet sites. We can use the FORE Native multicast to reach all of the sites using FORE equipment. For those sites using Cisco equipment we can set up an IP tunnel and allow them to redistribute the multicast traffic within their ATM network. Another option would be to designate a single multicast server and have it redistribute the traffic to all the sites.

The ring topology provides for two 20Mb/s bi-directional Virtual Paths at each site. This topology also allows for direct links between sites where existing applications are already in place (i.e., distance learning between UCSC and UC Extension, and BayLink between SJ Tech and MBA). The ring topology will also allow us to recover in the event that a single switch goes down. Other topologies using only the end nodes as routers are possible, such as a 2D mesh. This may be more useful if bandwidth hungry applications (BayLink and Distance Learning for example) require it. Obviously commercial adoption of ATM depends upon adoption and use of SVCs [2].

Distance Learning between MBA and MBARI

MBARI has two distance learning objectives as part of the CalREN ATM project: participate in distance learning originating from UCSC and the UCSC Extension; and create an operational distance learning capability between MBARI Moss Landing and the Monterey Bay Aquarium in Monterey. The MBA–MBARI distance-learning application suite will support the scientific training of LINKERs by MBARI scientists, and enable the sharing of scientific seminars between facilities. MBARI Moss Landing will be reconfigured into the regional ATM network in January 1996 to begin its participation in UCSC originated distance learning courses and REINAS regional weather briefings done by NPS.

MBA and MBARI have been working to define the requirements for the distance learning application between Monterey and Moss Landing. Our goal is to create a capability to send 2 bi-directional audio/video streams between MBARI auditorium or video laboratory and the MBA auditorium or Discovery Laboratory. In addition, we want to supplement these streams with an electronic white board to facilitate collaborative interaction between the source and receiver sites.

Current thinking proposes the use of our regional microwave infrastructure to support the audio/video transmission and to experiment with ATM capability as the medium for other collaborative resources such as multimedia data and white-board capability. Although this limits our initial outreach to Monterey Bay region schools, a link to a satellite uplink facility would allow a hybrid solution of satellite, cable, and digital network distribution of video, audio and data. If and when ATM technology stabilizes to provide an end-to-end, local and wide area network solution, we will be in position to move to that technical solution.

Research in Congestion Control for ATM networks

We have designed and analyzed congestion control algorithms for the support of *Available-Bit-Rate* (ABR) traffic in ATM networks. We have recently developed an efficient algorithm for rate allocation within the individual switches of an ATM network implementing the rate-based congestion control algorithm. The algorithm performs an allocation in O(1) time, allowing it to be applied to ATM switches supporting a large number of virtual circuits. Results from simulations using ATM sources show that the algorithm provides close to ideal throughput and converges to the max–min fair allocation rapidly when the available bandwidth or the individual requests change. We are also studying the behavior of the source algorithm in the rate-based congestion control framework to look at its effect on TCP behavior [6, 7, 8].

We are developing middleware for floor control, based on traditional concurrency control techniques, yet being able to adapt to a variety of media and their scheduling requirements. Mechanisms like floor control need to interface with established session control protocols and scale with the number of users, optimizing service delivery and supporting users transparently in their collaborative work. Creating floor control separate from applications has proven to be very challenging (CSpray has it built in). In order to assist the aquarium's scaling up of BayLink, floor control of a large number of users who can ask questions will be needed.

Conclusions

We described the nature of the BayNet ATM network, design, applications, and future plans. The CalREN funding expires this year, so all participants are actively evaluating the ATM use. Primary participants are looking for future funding and collaborations for broadband regional networks. The commercialization of ATM is occurring rapidly at this point, and the demand for switched virtual circuits, reliability, and manageability are being addressed and placed into the new standards. We have been working with nontraditional digital network applications to explore the future of broadband networks, and there remains much to be done and much to be learned.

More information

World-Wide Web URLs for the above sites:

```
http://www.cse.ucsc.edu/research/baynet-atm
http://csl.cse.ucsc.edu/reinas.html
http://www.stl.nps.navy.mil/~iirg
http://www.cse.ucsc.edu/research/slvg/
```

continued on next page

Monterey BayNet *(continued)*

Acknowledgements

We would like to thank all of the CalREN team at Pacific*Bell for their assistance, grant, and patience. We would like to especially thank Syd Leung and Ron Brown our CalREN managers. Thanks also goes to the Pacific*Bell technical assistance from Tom Soon. We cannot take sole credit for such an interdisciplinary and large scale project, and therefore thank all those who participated in BayNet. And of course thanks to the regional players who have made I3LA and BayNet-ATM a reality.

References

[1] D. Wiltzius, L. Berc, S. Devadhar, "BAGNet: Experiences with an ATM Metropolitan-area Network," *ConneXions,* Volume 10, No. 3, March 1996.

[2] P. Musich, "3Com's new products stiffen ATM backbone," *PCWeek,* Volume 13, No. 4, January 29, 1996, pp 1,100.

[3] H.-P. Dommel and J. J. Garcia-Luna-Aceves, "Floor Control for Networked Multimedia Applications," ACM SIGCOMM '95 Middleware Workshop, Cambridge, MA, August 28–September 1, 1995.

[4] Alex T. Pang, Craig M. Wittenbrink, and T. Goodman, "CSpray: a collaborative scientific visualization application," Proceedings IS&T/SPIE Symposium on Electronic Imaging: *Multimedia Computing and Networking,* Volume 2417, pp 317–326, February 1995.

[5] "Introduction to ATM Networks," 16-hour videotaped course by Prof. Anujan Varma of UCSC, available through the UCSC Extension.

[6] L. Kalampoukas and A. Varma, "Performance of TCP over Multi-Hop ATM Networks: A Comparative Study of ATM Layer Congestion Control Schemes," Technical Report UCSC-CRL-95-13, February 1995. A shortened version appeared in Proceedings of ICC '95, June 1995.

[7] L. Kalampoukas, A. Varma, and K. K. Ramakrishnan, "An Efficient Rate Allocation Algorithm for ATM Networks Providing Max–Min Fairness," Proceedings of the 6th IFIP International Conference on High-Performance Networking (HPN 95), September 1995, pp. 143–154.

[8] L. Kalampoukas and A. Varma, "Analysis of Source Policy in Rate-Controlled ATM Networks," Technical Report UCSC-CRL-95-??, September 1995.

[9] A. Pang, "Spray Rendering," *IEEE Computer Graphics and Applications,* Volume 14, No. 5, pp 57–63.

[10] M. Laubach, "IP Over ATM and the Construction of High-Speed Subnet Backbones," *ConneXions,* Volume 8, No. 7, July 1994.

[11] S. Minzer, "Broadband ISDN and Asynchronous Transfer Mode (ATM)," *IEEE Communications,* Volume 27, September 1989.

[12] M. de Prycker, *Asynchronous Transfer Mode—Solution for Broadband ISDN,* 2nd Ed, Alcatel Bell, 1993.

[13] R. Händel and M. Huber, *Integrated Broadband Networks,* Addison-Wesley, 1991.

[14] D. McDysan and D. Spohn, *ATM: Theory and Application,* McGraw Hill, 1995.

[15] "ATM Services, The Truth Behind the Hype," *Telephony,* September 1995.

[16] W. R. Stevens, *TCP/IP Illustrated, Volume 1: The Protocols,* Addison-Wesley, 1994.

[17] Laubach, Mark, "ATM for your internet—But When?" *ConneXions,* Volume 7, No. 9, September 1993.

[18] Atkinson, Ran, "Towards Real ATM Interoperability," *ConneXions,* Volume 7, No. 8, August 1993.

[19] Clapp, G. & Zeug, M., "Components of OSI: Asynchronous Transfer Mode," *ConneXions,* Volume 6, No. 4, April 1992.

CRAIG M. WITTENBRINK is a postdoctoral researcher at the University of California, Santa Cruz (UCSC) in Baskin Center for Computer Engineering and Computer Science. He develops architectures and algorithms for visualization and graphics, and wants the ATM testbeds to create the ultimate heterogeneous supercomputer. He worked as an Electrical Engineer digital designer developing parallel graphics computers for flight simulators at The Boeing Company. He received his B.S.('87) in Electrical Engineering and Computer Science from the University of Colorado, Boulder, and an M.S.('90) and Ph.D.('93) in Electrical Engineering from the University of Washington, Seattle. E-mail: `craig@cse.ucsc.edu`

ARUL ANANTHANARAYANAN is a member of the technical staff at the University of California, Santa Cruz, in the Baskin Center for Computer Engineering and Computer Science. He received his B.S. in Computer Science from San Diego State University in 1991, and is currently completing a thesis for an M.S. degree in Computer Science from San Diego State University. His interests include High Speed Networks and Internet routing protocols. E-mail: `arul@cse.ucsc.edu`

J. J. GARCIA-LUNA-ACEVES received the B.S. degree in electrical engineering from the Universidad Iberoamericana, Mexico City, Mexico, in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of Hawaii, Honolulu, HI, in 1980 and 1983, respectively. He is an Associate Professor of Computer Engineering at the University of California, Santa Cruz (UCSC). Prior to joining UCSC in 1993, he was a center Director at SRI International (SRI) in Menlo Park, California. He first joined SRI as an SRI International Fellow in 1982. Dr. Garcia-Luna has published more than 80 technical papers, has been guest editor for *IEEE COMPUTER* (1985), is an editor of the ACM *Multimedia Systems Journal,* and has chaired numerous networking conferences. E-mail: `jj@cse.ucsc.edu`

BRUCE GRITTON is Manager of Computer & Information Services at the Monterey Bay Aquarium Research Institute. Mr. Gritton has worked for twenty five years in the field of scientific and technical software systems in support of ocean science research, spacecraft control, and information systems. He is a primary facilitator for Monterey Bay collaborations, and current executive director of the Futures Network. E-mail: `grbr@mbari.org`

PATRICK MANTEY is the endowed Jack Baskin Professor of Computer Engineering and Chair of Computer Engineering at the University of California, Santa Cruz. He received his B.S., from the University of Notre Dame, M.S., from the University of Wisconsin-Madison, and Ph.D., from Stanford University. He has extensive computer industry experience, coming to UCSC from IBM, and has been and continues to be involved in the design of computer systems for a variety of applications in government and industry. E-mail: `mantey@cse.ucsc.edu`

Routing in a Multiprotocol over ATM Environment

by
Ross W. Callon, Bay Networks Incorporated,
Joel Halpern, Newbridge Networks,
John Drake, FORE Systems,
Hal Sandick, IBM Corporation,
Jason Jeffords, Cabletron Systems

The issue

Asynchronous Transfer Mode (ATM) offers high speed, low latency communications and enhanced Quality of Service (QoS) capabilities. However, for the foreseeable future a significant percentage of devices using an ATM network will do so indirectly, and will continue to be directly attached to "legacy" media (such as Ethernet and Token Ring), and will continue to make use of "legacy" internetwork layer protocols (such as IP, IPX, APPN, etc). This means that in order to effectively use ATM, there must be efficient methods available for operating multiple internetwork layer protocols over heterogeneous network environments made up of ATM switches, routers, frame switches, and hubs. This challenge is commonly referred to as the operation of "multiprotocol over ATM."

This article focuses on methods for routing in the multiprotocol over ATM environment. There are several aspects of this to consider, including:

- When to establish direct switched virtual channels (SVCs) between devices (i.e., routers or hosts) attached to an ATM network
- Routing of SVCs within the ATM subnetwork
- Routing of internetwork layer packets (e.g., routing of IP packets)

Switching and routing

Internetwork layer protocols are generally connectionless. This means that each packet is forwarded independently between routers without requiring maintenance of state information pertaining to any particular source/destination flow of data. In contrast, ATM is connection oriented. This means that ATM switches must obtain and store state information regarding the specific virtual circuit between the source and destination before forwarding user data. This is an important difference between internet layer protocols and ATM, and has a substantial effect on methods for running multiple internet layer protocols over ATM. This in particular impacts decisions regarding when to set up direct SVCs between devices, and when to use indirect paths via intermediate routers.

A fast router today can forward on the order of a million packets per second. ATM switches can establish on the order of hundreds or thousands of SVCs per second, with each SVC carrying from one to potentially millions of packets. It is therefore impractical to establish an SVC every time a packet is to be forwarded. Instead, it is necessary to establish a sufficient set of SVCs on an a priori basis, in order to ensure that an acceptable path will be available when a packet needs to be forwarded.

Given a heterogeneous network topology of routers and hosts attached to both legacy media and ATM, it is necessary to calculate internetwork layer (e.g., IP) routes consistently. The internetwork layer routing protocol must respond quickly to changes in the internetwork topology, including failure and recovery of equipment, installation of new equipment, and the opening and closing of SVCs. Many dynamic routing protocols are in common use which fulfill these requirements. OSPF and RIP are the most common multivendor standard protocols for use with IP.

An ATM subnetwork must to be able to route every SVC request via ATM switches from the calling to the called address. In a connection-oriented subnetwork, the path used by the SVC may stay the same for a potentially extended period of time. It is therefore important that the path be consistent, efficient, and loop free, even in the presence of transient changes in the network topology. The *Private Network-Network Interface* (PNNI) has been developed by the ATM Forum as a standard dynamic routing protocol designed specifically for use in routing SVCs across an ATM subnetwork.

The multiprotocol over ATM environment

Methods used for routing in a multiprotocol over ATM environment must be compatible with the equipment and capabilities that are expected to be needed in this environment. In many ways the multi-protocol over ATM environment is likely to be richer than traditional networking environments. Specific equipment and capabilities which must be accommodated include:

- ATM switches and links
- Legacy Networks (Ethernet, Token Ring, FDDI, point-to-point links, etc.)
- Virtual Networks
- Emulated LANs
- RFC 1577 logical IP subnets
- Virtual point-to-point links (e.g., permanent virtual channels (PVCs) using RFC 1483 encapsulation)
- Bridges (e.g., bridging legacy LANs and emulated LANs)
- Routers (enhanced to understand ATM-specific capabilities)
- "Virtual Routers" (route servers and non-routing edge devices)

ATM Switches and Links: ATM switches accept, process, and forward the ATM cells associated with a given SVC after that SVC has been established across an ATM subnetwork. Typically, this forwarding of 53 byte cells is done in hardware to provide extremely high speed, low latency communication. ATM links connect ATM switches to users and to other switches, and carry the ATM cells. Currently defined links range in speed from T1 (1.544 megabits per second) to OC-12 (622 Megabits per second). Additional links both at slower and faster speeds are being defined for use with ATM switching.

Legacy Networks: There are many existing technologies (including Ethernet, Token Ring, FDDI, Fibre Channel, HIPPI, and various speeds of point-to-point links), and new technologies are always being developed. Any attempt to address internetworking over ATM must allow for and interoperate with the operation of internetworking over any and all other technologies.

ATM Virtual Networks: ATM virtual networks can be thought of as ways to make one technology look like another technology. Emulated LANs, using the ATM Forum LANE specifications, for example, allow a collection of stations to treat an ATM fabric as if, in addition to being ATM, it were an Ethernet or Token Ring segment.

RFC 1577 Logical IP subnets are another approach to making ATM look like something familiar. The station treats ATM as an unspecified subnetwork with "conventional" connectivity properties such as it would see on Ethernet, Token Ring, or FDDI. It does so at the IP layer, without introducing 802.1 addressing or framing.

continued on next page

The Multiprotocol ATM Environment *(continued)*

Alternatively, one could use PVCs and a well defined encapsulation (as, for example given in RFC 1483) to provide the equivalent of what the internetwork layer would expect from a collection of point-to-point links.

Bridges: Bridges are an example of one method for interconnecting ATM with legacy subnetworks. In this context we mean any device which performs forwarding based strictly on the MAC address in a received frame. This includes traditional Ethernet and Token Ring bridges, hubs, switching hubs, and the many variations thereon.

Routers: Routers are the current method for providing internetworking services between heterogeneous subnetworks. Routers may be defined to be devices which perform forwarding based on the internetwork address (and other internetwork layer information) in a received packet. In most cases routers will also participate in the operation of one or more internetwork level routing protocols (such as OSPF, RIP, or I-PNNI). These devices, or variations thereon, are the key to providing multiprotocol interconnection and operation for ATM.

Virtual Routers: Virtual routers are a new wrinkle in the routing environment. Traditional routers have two closely related functions. The first is the exchange of internetwork layer routing information with other routers to determine the internetwork layer topology. The second is the forwarding of internetwork packets.

In a virtual router, these two functions are logically (and possibly physically) separated. If the functions are separated, one device, the *route server,* exchanges internetwork layer routing information with other routers (and route servers); Another device, the *edge device forwarder,* performs the forwarding of internetwork packets. The two types of devices together provide the functionality of a traditional router.

Options for routing

Operation of internetwork layer protocols over ATM is still relatively new and has primarily used ATM to emulate point-to-point links and simple LANs. This is the most straightforward way to make use of ATM in existing networks, but does not make full use of the capabilities of ATM. The best methods for running internet level protocol over ATM is still very much an area for ongoing research, experimentation, and standardization.

The remainder of this article explores three options for routing in this environment, which are:

- Layered Routing
- PNNI Augmented Routing
- Integrated PNNI

With *Layered Routing,* the routing protocol used within the ATM subnetwork is completely separate from the routing protocol used between routers, and routers do not have any knowledge of the internal structure of the ATM subnetwork.

PNNI Augmented Routing also uses separate routing protocols for internetwork layer routing and for the ATM subnetwork. However, routers (and route servers) with ATM interfaces also participate in the operation of the ATM routing protocol (e.g., PNNI). This allows routers to locate other routers across the ATM subnetwork, thus providing an auto-configuration capability.

In addition, this allows those routers with ATM interfaces to understand the internal structure of the ATM subnetwork. This is helpful in order to coordinate the management of SVCs across the ATM subnetwork.

Integrated PNNI (I-PNNI) makes use of a single routing protocol (based on the PNNI standard) for routing between routers, route servers, and ATM switches. This allows a single routing protocol to be used for an entire routing domain. This therefore allows end-to-end routing to be based on the true network topology, provides for auto-configuration of SVCs (as above), and requires that only one routing protocol be managed and configured.

Each of these possible methods will be appropriate for some environments. Each is described in more detail below.

A simple example network is shown in Figure 1. This figure may be used to illustrate operation of the three options described above. In the figure there are ten routers, numbered R1 through R10, as well as three ATM switches A, B, and C. A routing protocol is required to determine how to route SVCs between the ATM switches. Similarly, given some set of SVCs across the ATM subnetwork connecting the routers, a routing protocol is required for routing between the routers. Finally, some method is required for maintaining the set of SVCs between those routers with ATM interfaces.

Route servers and edge device forwarders are not illustrated in Figure 1. This is in order to simplify the example, and is not meant to minimize the importance of route servers and edge device forwarders. Similarly, the protocols must support routing to virtual networks which may be distributed amongst multiple routers and/or edge device forwarders.

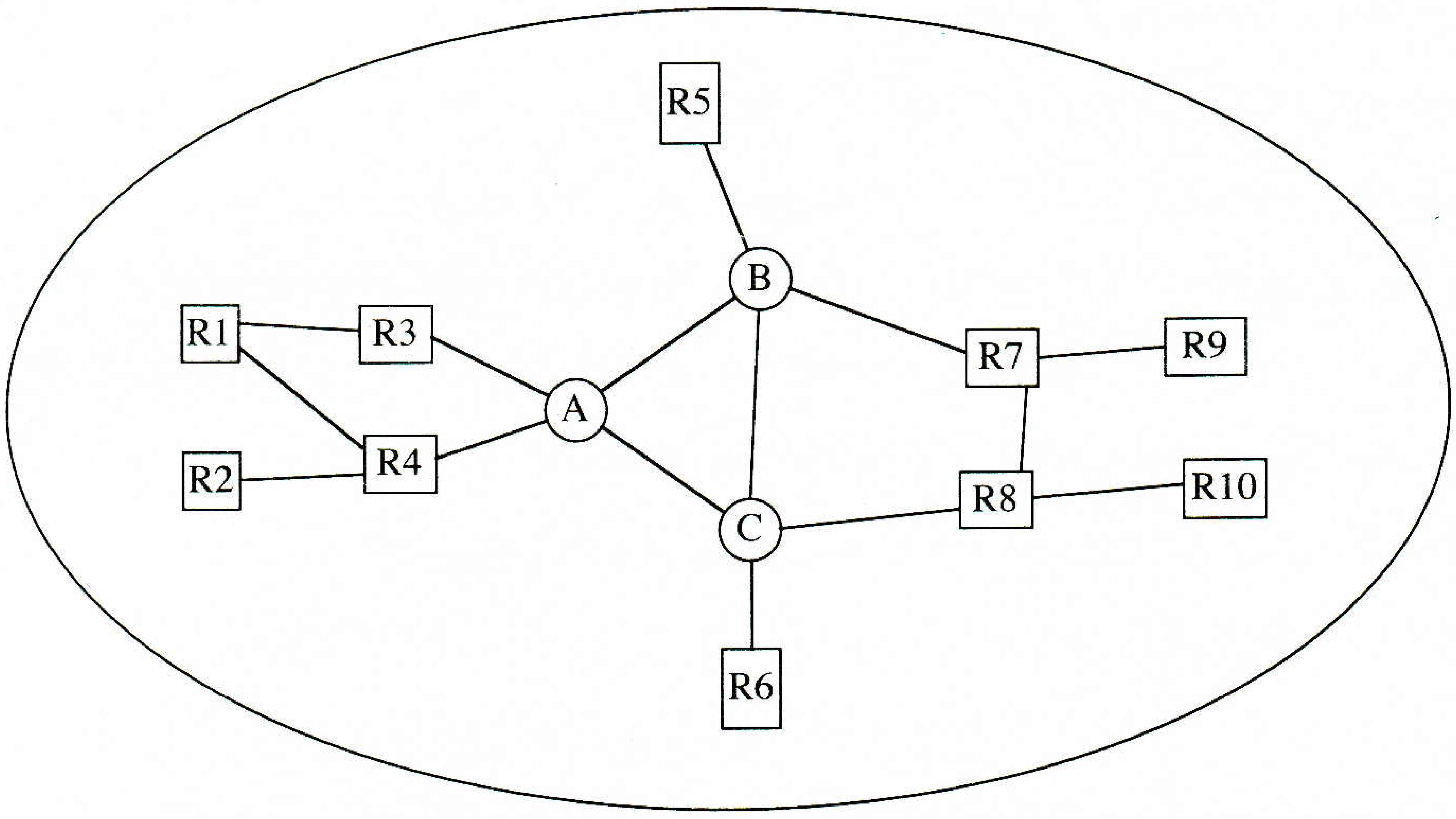

Figure 1: Example Network with Routers and ATM Switches

Layered Routing

Layered Routing is simply the direct transference of traditional routing techniques to the ATM environment.

RFC 1483 Permanent Virtual Circuits: RFC 1483 defines a simple encapsulation which may be used to run internet layer protocols over ATM. This may be used along with Permanent Virtual Circuits (PVCs) to emulate point to point links. When used in this manner ATM provides high speed point-to-point interconnect as well as a relatively easy means to reconfigure the network when necessary.

continued on next page

The Multiprotocol ATM Environment *(continued)*

LAN Emulation: ATM Forum standard LAN emulation allows a set of ATM-attached equipment to interoperate as if it were attached to a simple LAN segment (Ethernet or Token Ring). If a set of emulated LANs are established such that all routers are attached to one or more emulated LANs (and possibly to actual legacy subnetworks), then the routers behave exactly as if they were connected to Ethernet or Token Ring segments. If the set of emulated LANs is created such that there is a path, made up of routers and emulated LAN segments, between any two points on the edge of the ATM network, then all stations can communicate.

Logical IP subnets: The simplest next step is to observe that when they are using internetworking protocols, the routers and hosts attached to ATM do not need the MAC layer services. So one can create a logical subnet structure, while omitting the mechanisms for pretending to be a legacy segment. RFC 1577 describes how this is done for IP.

Clearly, having built logical IP subnets, one can again use routers to connect them. This, coupled with the Multicast Address Resolution and Multicast Server work being developed, provides a full set of internetwork layer capabilities, again for IP.

Remote Address Resolution: By adding one additional piece of capability to the logical IP subnet, one gets a new service that makes better use of the underlying ATM fabric connectivity.

First, a bit of explanation. Currently, an IP host makes a decision as to whether a destination is local or remote. Based on that decision, it either sends the packets for that destination to a router, if it is remote, or performs address resolution to find the destination, if it is local. On legacy subnetworks, the local/remote decision is done by comparing the destination address with the local host address, using a mask to decide what bits are significant (this is frequently referred to as "mask and match"). This reflects the address assignment rules for legacy subnetworks, where stations on the same subnetwork are assigned addresses from the same set of numbers (subnet).

However, on an ATM network the entire world is potentially local. Therefore, a mask and match determination is not necessarily appropriate. Instead, the host (or router) can make a decision based on traffic (and quality of service). For short transactions (such as name resolution), a default path, using the overlay described under the Logical IP subnets, but not even doing local resolution, is quite sufficient.

At the same time, for a longer transaction it may be desirable to establish an SVC directly to the destination. Therefore, a query protocol has been developed by the IETF for this purpose. The NHRP query (and response) uses the routing infrastructure to provide internetwork to ATM address resolution.

With NHRP, hosts or routers can run traditional routing, and also get access to direct ATM connectivity (direct SVCs) when desired. This is also important in the context of the virtual router, where it may be highly desirable that traffic switch to direct SVCs for any persistent transactions.

There are, of course, additional issues.

Firstly, the routers serving a given logical IP subnet must talk to each other. Therefore, there must be either configured or dynamic methods for them to discover each other. The most obvious technique is to use a protocol like OSPF, which uses multicast, and use the multicast mechanisms to find and pass the information among the routers. This assumes bootstrapping via a configuration service using a well known ATM address. Manual configuration is also possible. Better techniques are explored below.

Secondly, there is the question of what connectivity the routers should advertise to other legacy-attached routers. Using the logical IP subnet approach, the routers advertise connectivity to the ATM subnetwork, and connectivity to each other may be inferred from this. While sufficient, this produces a limited view of the world.

When advertising connectivity, there is one important caveat. Reachability based on SVCs created by routers as a result of query (NHRP) responses must not be advertised directly into the routing protocols. The routing protocols provide the stability that allows these SVCs to be established, and advertising them could undermine this stability.

PNNI Augmented Routing

PNNI Augmented Routing consists of running legacy internetwork layer routing protocols (such as OSPF, RIP, or NLSP) over legacy media and also over ATM, but in addition, having routers with ATM interfaces also run the PNNI protocols. Routers with ATM interfaces advertise their status into PNNI, and use information about other routers gained from PNNI to bootstrap and maintain the SVCs between routers. This simplifies the management of the legacy routing protocol over ATM, and improves the maintenance of SVCs across the ATM subnetwork used for routing.

When this is done, all routers run one or more internetwork layer routing protocols, and use the results for forwarding of internetwork packets. All ATM switches run standard PNNI routing (i.e., PNNI Phase 1). Those routers with ATM interfaces in addition run standard PNNI routing (i.e., participate in the operation of the ATM PNNI routing protocol). The information obtained from the PNNI routing protocol allows these routers to locate other routers on the ATM subnetwork, and provides these routers with detailed internal knowledge of the state of the ATM subnetwork. This eliminates the need for manual configuration of virtual circuits across the ATM subnetwork, thereby simplifying the initialization and network management for these routers. This also allows more efficient maintenance of SVCs between these routers across the ATM subnetwork.

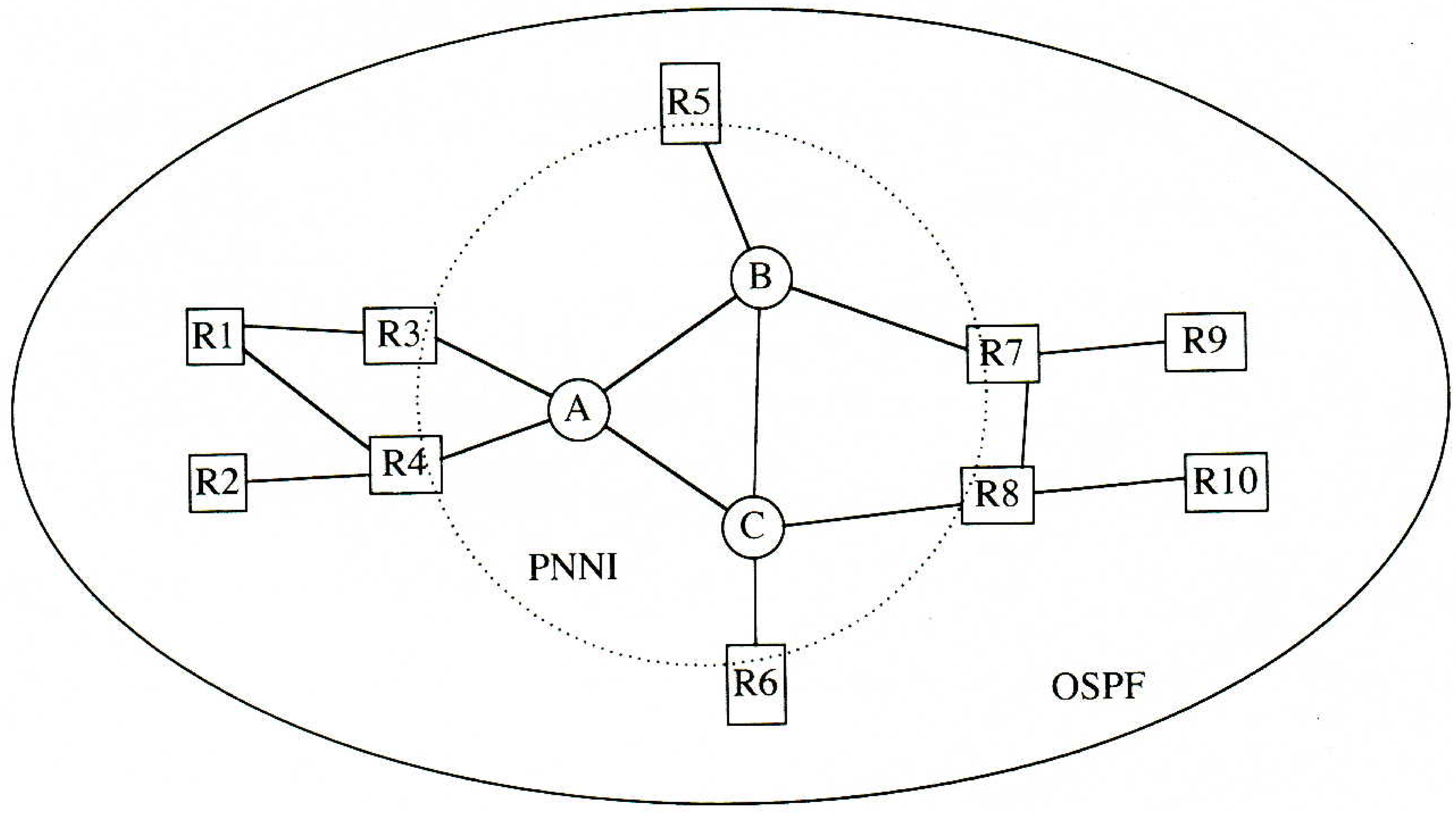

Figure 2: PNNI Augmented Routing

continued on next page

The Multiprotocol ATM Environment *(continued)*

In our example let's suppose that IP is the internet layer protocol of interest, and that OSPF is the preferred internetwork layer routing protocol. In this case all ten routers participate in OSPF. The ATM switches and those routers with ATM interfaces (routers R3 through R8) participate in the operation of PNNI. The interfaces between these routers and ATM switches are therefore PNNI interfaces, and these routers understand the combined internetwork layer and ATM subnetwork topologies as illustrated in Figure 2 on the previous page.

PNNI has been defined specifically to be extensible. Therefore, router specific information can be advertised in PNNI packets using Type-Length-Value encoding that ATM switches will know to ignore. Also, a router would advertise itself as a "transit restricted" ATM switch. This means that is is acceptable to route an SVC to that router if the SVC ends at that router, but it is not acceptable to route an SVC to that router if the SVC is to transit that router. This allows the routers implementing PNNI to be fully compatible with ATM switches that implement only PNNI Phase 1. The ATM switches are not required to have any knowledge of internetwork layer routing.

Route advertisements

In running an internetwork layer routing protocol, it is necessary to characterize/represent the capabilities of the ATM subnetwork. Even if there are no SVCs currently established across the ATM subnetwork (or no SVCs to some particular destinations), it is still necessary to advertise into the router network that it is possible to forward packets across the ATM subnetwork. For example, this allows routers which are not directly connected to the ATM subnetwork (such as routers R1 and R2 in Figure 1) to route packets across the ATM subnetwork to destinations across the ATM subnetwork.

There are several ways that this may be accomplished. One method is for the routers with ATM interfaces to establish a set of "default SVCs" between themselves immediately upon network initialization. The establishment of this set of SVCs is based on the knowledge gained from running PNNI, and allows creation of a set of SVCs which span the ATM subnetwork. Thus for any two routers, X and Y, which are participating in PNNI, and which are (for example) in the same OSPF area, there might not necessarily be a direct SVC from X to Y. However there will be some path from X to Y using only other routers in the same OSPF area as intermediate hops.

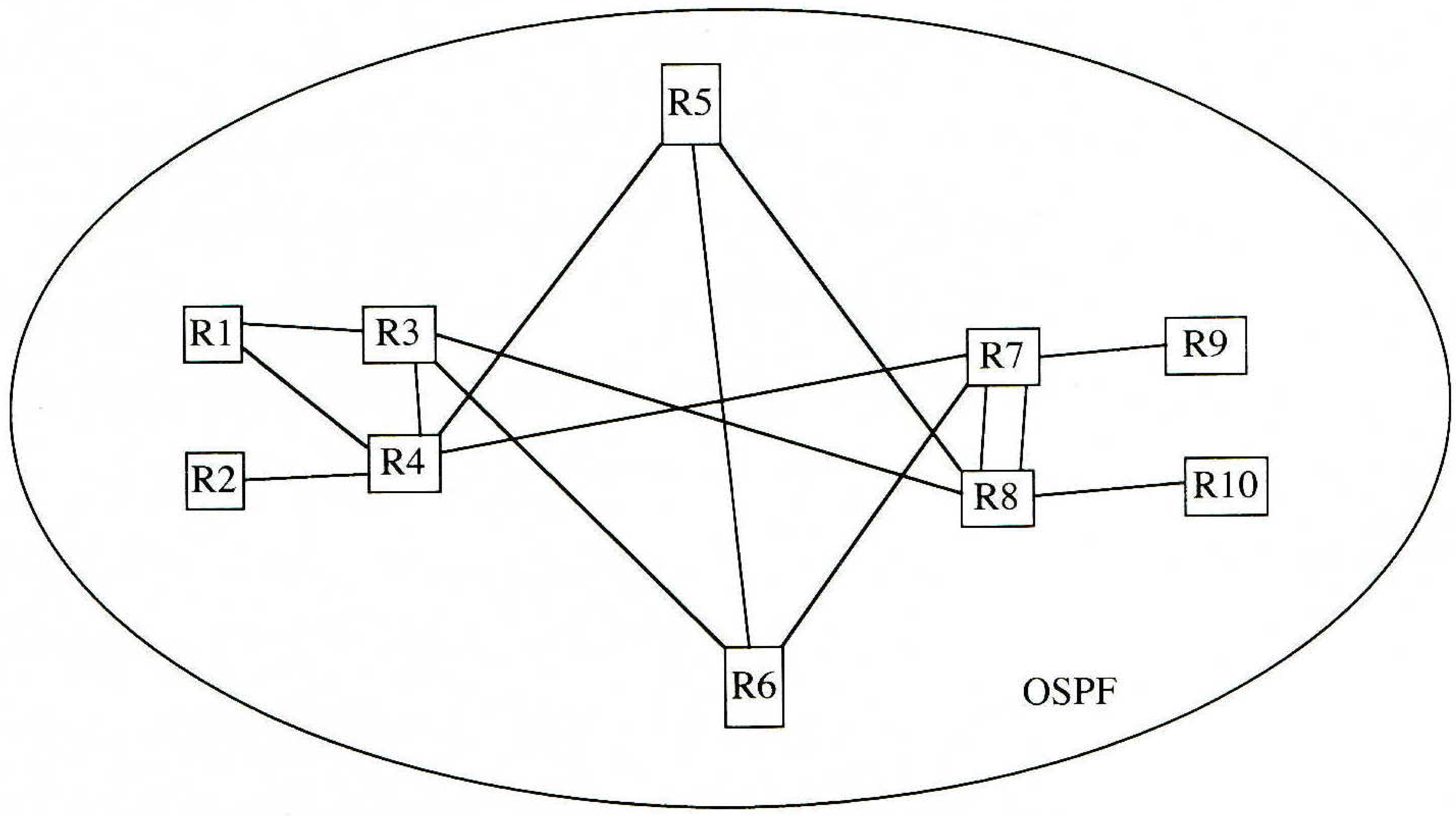

Figure 3: Logical Router Topology with Default Set of SVCs

As an example, for small networks the routers with ATM interfaces may choose to establish a full mesh of SVCs between themselves. For larger networks, a partial mesh of SVCs would be chosen. The choice would be made automatically by the routers, based upon the size of the network, among other things. Let's suppose that in our example each router chooses to establish an SVC to the next higher numbered router (with the number space treated as circular, so that the highest numbered router opens an SVC to the lowest numbered router), and in addition to the third higher numbered router. This would result in the logical OSPF topology illustrated in Figure 3.

With this method, the set of default SVCs are explicitly advertised into the router network with OSPF, for example. OSPF would operate over these SVCs as if they were normal point-to-point links. If additional routers joined the network, then additional SVCs would be established automatically, without the need for manual configuration of each SVC. Note that this does not preclude user configuration of additional PVCs or SVCs.

With larger networks, it will not be feasible to establish SVCs between every possible pair of routers. This implies that the path between any two routers may require multiple hops across the ATM subnetwork. When there is a small amount of traffic taking extra hops, the extra hops may be preferable to setting up additional SVCs. When there is a substantial amount of traffic taking extra hops the routers running PNNI will have sufficient information to determine whether it is worthwhile to establish a direct SVC. If there is a topology change, these routers would have sufficient information to determine whether the direct SVCs are still appropriate to reach any particular destination.

Routers running PNNI participate as regular PNNI nodes, exchange PNNI Hellos with neighboring ATM switches, and broadcast regular PNNI PTSPs throughout the peer group. These routers make use of standard PNNI packets to announce normal PNNI information such as links to neighboring ATM switches, metrics on these links, and the node ID and ATM address of the router. In addition, the router uses the extensibility features of PNNI to announce (in a way that normal ATM switches will ignore) router specific information such as the internetwork layer protocol suite supported (e.g., IP), a router ID, the internetwork layer routing protocol used (e.g., OSPF or RIP), and routing-protocol specific information (such as the OSPF area).

PNNI augmented routing may be thought of as an optimization of Layered Routing. It allows routers (and route servers) to find each other across the ATM subnetwork, and thereby facilitates auto-configuration of SVCs between routers (and route servers). It also ensures that a reasonable set of default SVCs is established automatically, as part of network initialization. This reduces the need to buffer or discard internetwork packets while waiting for queries to be answered or SVCs to be established. It provides automatic adjustment of the SVCs between routers as routers join or leave a PNNI peer group and it provides automatic adjustment of the SVCs based on transient features such as traffic load.

PNNI augmented routing allows existing internetwork routing protocols to operate essentially unchanged over a combination of existing networks and ATM subnetworks. It therefore provides a straightforward incremental addition to existing router networks.

continued on next page

The Multiprotocol ATM Environment *(continued)*

PNNI augmented routing is also simple enough that it would be straightforward to extend it to support other internet level protocol suites (such as IPX, DECnet, AppleTalk, etc).

PNNI augmented routing may be extended for operation in a multi-level hierarchy. There are two cases of this: (i) operation in a PNNI hierarchy; (ii) Operation where the internetwork level protocol (e.g., OSPF) is itself hierarchical.

Consider operation of a single OSPF area which spans multiple PNNI peer groups, in a common higher level PNNI parent peer group. In this case, those routers in the same OSPF area which are also in the same PNNI peer group contact each other and establish a set of default SVCs as described above. The router with the highest router ID sets the attribute bits in its PNNI advertisement to indicate that this advertisement should be advertised into the parent peer group. This allows routers in other PNNI peer groups (within the same parent peer group) to discover the existence of this router. This in turn allows a set of default SVCs to be established between routers in different PNNI peer groups, allowing the OSPF area to span multiple PNNI peer groups.

Separate OSPF areas may operate independently, each using PNNI to coordinate its operation over an ATM subnetwork. Area border routers may participate in multiple areas. This allows multi-level OSPF to operate in essentially the same manner as in any other multilevel OSPF routing domain. This does imply that in some cases a packet may take multiple hops across the ATM subnetwork (one from a router within an area to an area border router, plus another hop in the other area served by the area border router). When necessary, this in turn may be shortcutted using a query response protocol such as NHRP.

Integrated PNNI

Integrated PNNI (I-PNNI) involves use of a single protocol (PNNI) for routing of one or more internetwork layer protocols as well as ATM. In order for this to adequately support SVC routing in the ATM network, use of integrated routing to support internetwork layer and ATM subnetwork routing is in practice limited to PNNI as the preferred routing protocol.

I-PNNI allows a single instance of PNNI routing to be run by routers (including route servers) and ATM switches. PNNI therefore is run between a switch and neighboring switches, between a switch and neighboring routers, and between routers. Thus all ten routers in Figure 1 (including routers which have no ATM interface) and all three switches run the normal PNNI routing protocol, and all appear as "nodes" in the topology database. The interface between a router and a switch (such as the links between switch A and routers R3 and R4 in Figure 1) runs the PNNI protocol and operates as a normal PNNI interface. PNNI Hellos and PTSEs/PTSPs are exchanged between all nodes (including routers, route servers, and switches) in the network. Similarly, all nodes and links are advertised using normal PNNI metrics. Thus the router-to-router links are advertised using the same metrics as are used for inter-switch links

The basic protocol operation is unchanged from PNNI (which in turn is in many ways typical of link state and topology state routing protocols). Each node exchanges Hello packets with its immediate neighbors, in order to reliably determine its local topology.

Topology information

Each node then advertises its local topology in "Topology State Packets" which are reliably flooded to all nodes in the network (or in the same peer group where hierarchical routing is used). This allows each node to know the full topology of the network (or group), which in turn allows each node to calculate routes to other nodes in the group.

I-PNNI is fully compatible with ATM switches which run PNNI Phase 1. ATM switches are not required to have any knowledge whatsoever about the internet level. This makes use of the extensibility features specified in the ATM PNNI Phase 1 standard (as was discussed above).

I-PNNI may run over a combination of ATM subnetworks, and legacy links and subnetworks. This means that I-PNNI may be run over broadcast LANs (such as Ethernets). Efficient operation over broadcast LANs is accomplished through use of the same pseudo-node and designated router capability that is used for OSPF and IS–IS.

In I-PNNI, both reachability to ATM systems and reachability to internet systems (such as IP subnets) is advertised into PNNI. However, different types of reachability are advertised independently. Thus, for example, reachability to ATM systems is advertised using the format defined in the PNNI Phase 1 standard. Reachability to IP subnets is advertised independently. Thus a router with an ATM interface may advertise "My ATM address is xxxxxxxxxx," and also separately advertise "I can reach IP subnet yyyy with mask zzzz."

Routers which are not attached to ATM subnetworks may nonetheless participate in I-PNNI. A router which is not ATM attached would advertise itself as a PNNI node (using normal PNNI packet formats), would advertise links to neighbors, and would advertise its IP addresses and reachability to IP subnets.

Virtual Subnet support

I-PNNI is intended for use in a multiprotocol over ATM environment, and therefore includes explicit support for features such as virtual subnets that people expect to use in this environment. An example of this is illustrated in Figure 4.

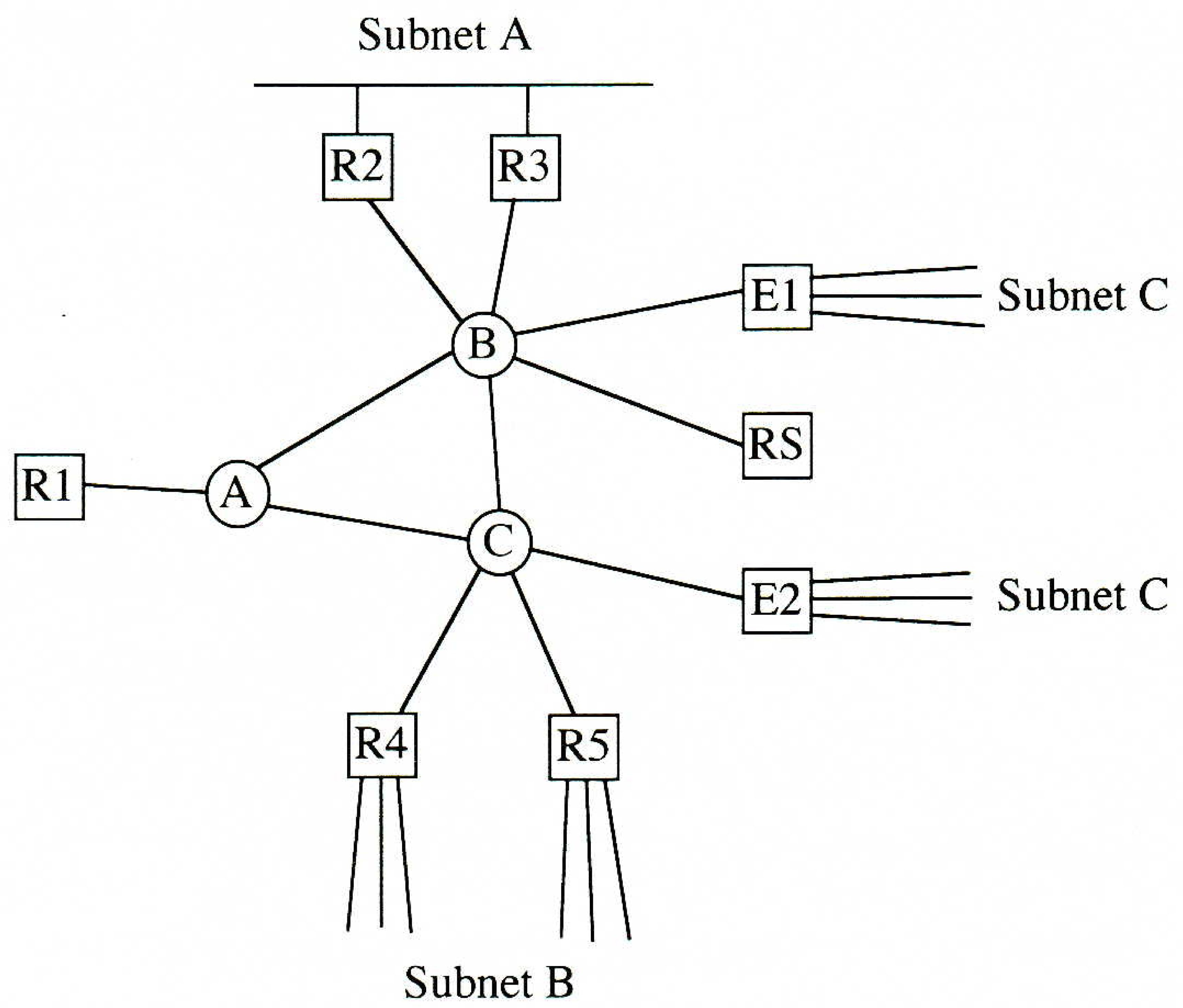

Figure 4: Special Requirements for MPOA Environment

continued on next page

The Multiprotocol ATM Environment *(continued)*

Figure 4 illustrates a multiprotocol over ATM network. Routers R1, R2, and R3 are normal routers, which implement I-PNNI, NHRP, and normal IP packet forwarding. Routers R4 and R5 in addition implement the ability to support virtual networks divided between multiple routers. Thus subnet B is partly reachable via R4, and partly reachable via R5. Finally, route server RS and Edge Devices E1 and E2 implement a distributed router, which supports virtual subnet C.

Suppose that there are three packets at R1, one each destined for subnets A, B, and C. With traditional routing protocols, there is no way to distinguish these three cases. This will require either conspicuous use of NHRP queries, or potentially inefficient routing (implying extra hops across the ATM subnetwork). In contrast, I-PNNI allows nodes to distinguish between "direct reachability," versus "query reachability." Direct reachability occurs when the entire subnet is reachable by the node directly, and there is no need to use a query (for example when forwarding packets for subnet A). Query reachability implies that while the node can forward packets to the subnet, in general if there are a lot of packets to a destination on the subnet it may be more efficient to query first (for example this occurs when forwarding packets to subnet B or C). The ability to distinguish between these different cases is a very simple addition to the routing protocol, but is an example of how I-PNNI is designed specifically for the multiprotocol over ATM environment, and represents a capability which is not available with other routing protocols.

Routing hierarchies

I-PNNI can operate in a multilevel hierarchy, and makes use of the same flexible hierarchy mechanisms which are defined for PNNI Phase 1.

One example of the use of this is illustrated in Figure 5. Figure 5 illustrates the physical interconnection between devices. Here a large corporate network makes use of a high speed ATM backbone, plus multiple local campus area networks. Some of the campus area networks may use only ATM switches, some may use only routers, and others may use routers interconnected via a local ATM backbone. Note that in the picture the illustrated routers may in fact consist of a combination of traditional routers, "level 3 switches" (router/hub/frame switch combinations), and route server/non-routing edge device combinations. Also note that while only 14 routers and 16 ATM switches are illustrated, the figure is intended to illustrate a much larger network. In fact a network with only 30 nodes would not require hierarchical routing, but a network large enough to require hierarchical routing is impractical to illustrate in a single figure.

The most straightforward way to handle this with the PNNI hierarchy (assuming that hierarchical routing is required) is to have the backbone operate at a particular level of the hierarchy, and to have each local campus network operate as a separate peer group at a lower level. The Peer Group Leader in each peer group selects the backbone as the appropriate higher level parent peer group. This implies that each of the campus networks is represented as a single node in the higher level backbone peer group. For details of how the PNNI hierarchy would be applied in this case, see the ATM Forum PNNI Phase 1 specification.

A large hierarchical network of this sort is a very straightforward fit to Integrated PNNI and the PNNI hierarchy mechanisms. This network would be more difficult to represent using pure Layered Routing. Assuming that the total network size is on the order of 100 nodes or larger, the ATM network is too large to efficiently represent the interconnection between routers as a simple emulated LAN. This implies that a spanning set of VCs would need to be configured between multiple routers, and that a multilevel router hierarchy would need to be configured. PNNI augmented routing would provide some help here, but would not eliminate the need to configure a hierarchy of routers. For example, if OSPF were used, then some routers would need to be configured to participate in the backbone OSPF area (area 0), and the router hierarchy would need to be configured in addition to the ATM hierarchy.

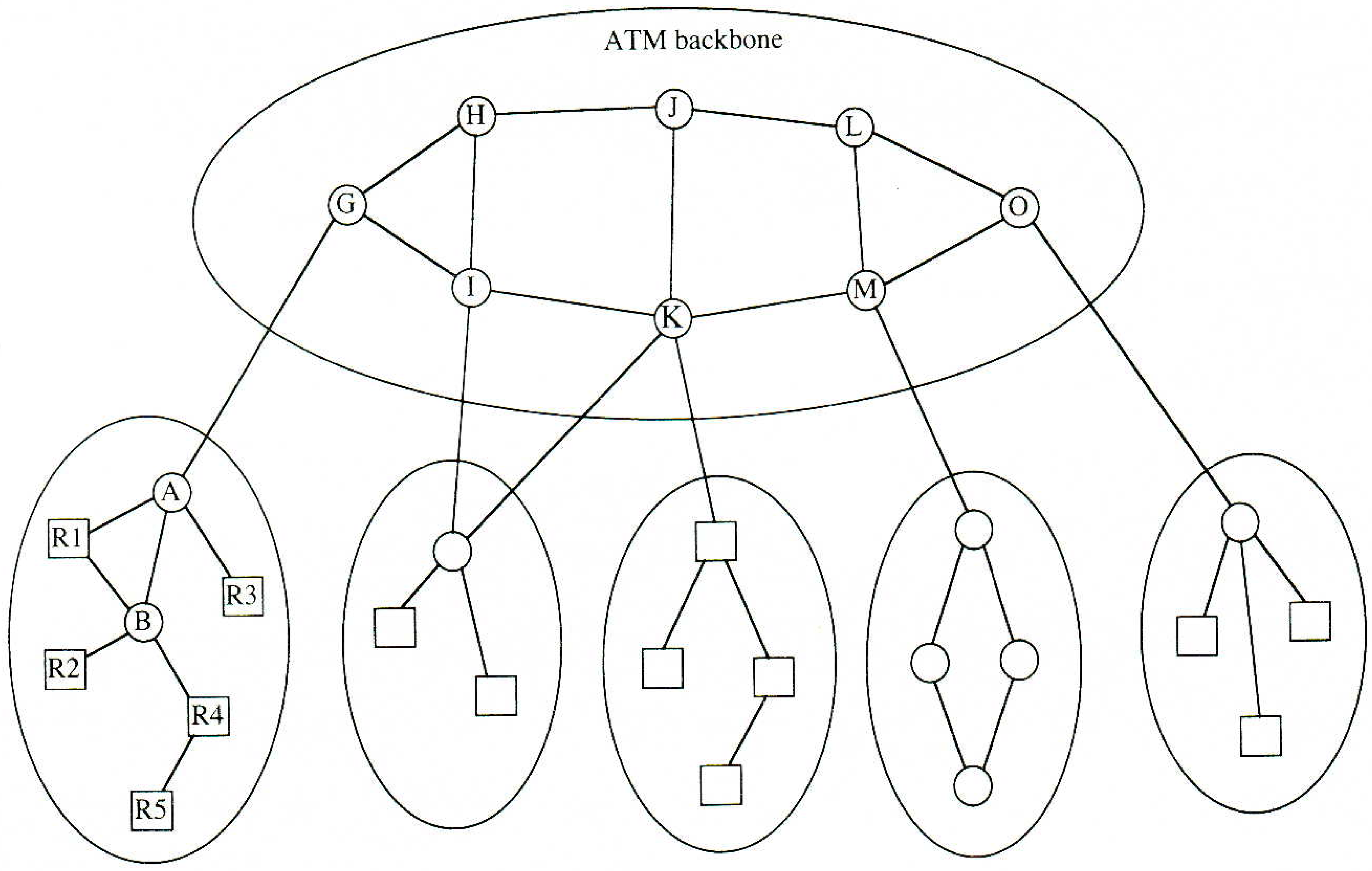

Figure 5: Example Network With High Speed ATM Backbone

QoS support

I-PNNI extends the capabilities of PNNI to internet level routing. For example, this includes QoS routing capabilities and flexible hierarchy mechanisms. I-PNNI also allows end-to-end routing through a combination of routers and ATM switches based on the true combined network topology. In some topologies (particularly where there are choices in the places where packets can enter and exit the ATM subnetwork) this will result in better end-to-end routes. This allows end-to-end QoS routing in a heterogeneous network environment.

I-PNNI also allows routers to locate each other across the ATM subnetwork, and aids in the coordination of SVCs across the ATM subnetwork. I-PNNI allows the user to run only one protocol for both IP and ATM (given that the user will need to use PNNI for ATM routing, this avoids the use of a second routing protocol to support IP).

Use of Integrated PNNI requires that ATM switches need to store information broadcast by routers within the same peer group, even though the switches won't actually route via the routers (except for SVCs ending at a router). This implies that Integrated PNNI scales as if the routers and switches were all switches. However, switches do not need to interpret the router information, and do not need any capability other than that specified in the ATM PNNI Phase 1 standard.

continued on next page

The Multiprotocol ATM Environment *(continued)*

Other issues

It is not possible in a short article such as this to give a complete detailed description of all of the methods for routing in a multiprotocol over ATM environment. This section briefly presents some of the other issues which are important for each of these proposals, but which were excluded from the discussions above in order to shorten the length of this article.

The biggest part that has been left out is the detail of protocol operation. Details which are being specified elsewhere include information on how reachability is advertised into the protocol, how I-PNNI relates to NHRP (these may be used in combination in many circumstances), and the detailed operation of the hierarchy for both PNNI augmented routing and I-PNNI.

Given that I-PNNI is intended for use as the routing protocol for IP (and other protocols) within a routing domain, it is necessary to be able to advertise external routes into the protocol. Similarly, the details of BGP/I-PNNI interaction need to be specified.

A critical issue in any communications network is the manner in which the network responds to changes in the network, such as failure and recovery of network equipment, transient congestion, or introduction of new equipment into the network. The details of network operation under changing conditions are very important, but are in some cases complex and have therefore not been discussed in detail in this article.

Work is needed to address the criteria routers should use to determine when to set up and release short cut SVCs. Work is also needed to determine how private virtual networks can be supported.

Also, methods are needed to allow graceful migration between Layered Routing, PNNI Augmented Routing, and I-PNNI. These and other related topics are the subjects of ongoing work within the ATM Forum.

Conclusion

There are a variety of methods available for routing in a multiprotocol over ATM environment.

For a large router network which is currently running an existing routing protocol such as OSPF, there is likely to be a reluctance to change the routing protocol in use. However, in networks which include a large ATM subnet it will be desirable to use PNNI over the ATM subnet, and the change from PNNI to I-PNNI is relatively minor.

Thus for initial deployment in large router networks with a very small amount of ATM, Layered Routing may be preferred. As the size of the ATM subnet grows PNNI will be needed in the ATM subnet and PNNI Augmented Routing becomes preferable. Finally, as the size of the ATM subnet continues to grow and as the user gets more familiar with the operation of PNNI in the ATM subnet, for environments in which end to end QoS routing is needed, or for new ATM-centric networks, Integrated PNNI becomes the logical choice.

The interaction between I-PNNI and OSPF (or other legacy routing protocols) will also depend upon the topology of the network. For example, if there is a large existing router network interconnected in a limited number of places to a newer ATM-centric network (with routers or other edge devices providing IP service across the ATM-centric network), then it may make sense to run OSPF in the router-centric part of the network and I-PNNI in the ATM-centric part of the network.

This implies that the choice between Layered Routing, PNNI-Augmented Routing, and Integrated PNNI may depend upon the specific network. The three methods mentioned here (Layered Routing, PNNI Augmented Routing, and Integrated PNNI) are each desirable for use in some environments under some conditions. Standards work is ongoing to ensure that all three methods for routing in a multiprotocol over ATM environment are available to users, including users with multivendor networks.

References

[1] PNNI Subworking group, "PNNI Draft Specification," edited by R. Cherukuri and D. Dykeman, ATM Forum 94-0471R15, January 1996.

[2] D. Katz, D. Piscitello, B. Cole, James Luciani, "NBMA Next Hop Resolution Protocol (NHRP)," Internet Draft, December 1995.

[3] LAN Emulation Subworking Group, "LAN Emulation over ATM, Version 1.0," edited by B. Ellington, AF-LANE-0021.000, January 1995.

[4] MPOA Subworking Group, "Baseline Text for MPOA," edited by C. Brown, ATM Forum 95-0824R5, January 1996.

[5] M. Laubach, "Classical IP and ARP over ATM," RFC 1577, January 1994.

[6] J. Moy, "OSPF Version 2," RFC 1583, March 1994.

[7] Y. Rekhter, D. Kandlur, "Local/Remote Forwarding Decision in Switched Data Link Subnetworks," Internet Draft, December 1995.

[8] R. Callon, "Integrated PNNI for Multi-Protocol Routing," ATM Forum 94-0789, September 1994.

[9] L. Kleinrock, K. Farouk, "Hierarchical Routing for Large Networks," *Computer Networks,* Volume 1 No. 3, January 1977.

[10] J. McQuillan, I. Richer, E. Rosen, "ARPANET Routing Algorithm Improvements, Volume 1," BBN Report 3803, April 1978.

[11] Laubach, Mark, "ATM for your internet—But When?" *ConneXions,* Volume 7, No. 9, September 1993.

[12] Atkinson, Ran, "Towards Real ATM Interoperability," *ConneXions,* Volume 7, No. 8, August 1993.

[13] Laubach, M., "IP Over ATM and the Construction of High-Speed Subnet Backbones," *ConneXions,* Volume 8, No. 7, July 1994.

[14] Clapp, G. & Zeug, M., "Components of OSI: Asynchronous Transfer Mode," *ConneXions,* Volume 6, No. 4, April 1992.

[15] Rybczynski, Tony, "The ATM Enterprise Network Switch: Enabler of Network Transformation," *ConneXions,* Volume 9, No. 11, November 1995.

[16] Tsuchiya, P. (now Francis, P.), "Components of OSI: IS–IS Intra-Domain Routing," *ConneXions,* Volume 3, No. 8, August 1989.

[17] Hedrick, C., "Routing Information Protocol," RFC 1058, June 1988.

[18] *ConneXions,* Volume 3, No. 8, August 1989, Special Issue: Internet Routing.

continued on next page

The Multiprotocol ATM Environment *(continued)*

ROSS CALLON is a consulting engineer with Bay Networks Incorporated in Billerica, Massachusetts, and has more than 15 years experience in data communications. He has been a major contributor to the ATM Forum PNNI routing standard, and is also involved in the Multiprotocol over ATM subworking group. He is the original author of the Integrated PNNI proposal. He has also been active in the IETF, and is a former co-chair of the IPng working group.
E-mail: `rcallon@baynetworks.com`

JOEL M. HALPERN is a principal engineer with Newbridge Networks Inc, in Herndon, Virginia. He has more than 10 years experience with data communications. As the Routing Area Director in the IETF, he oversees and assits ongoing developments in IP routing technology. Within the ATM Forum, Joel is one of the major contributors to the PNNI (Private Network-Network Interface) routing work and to the MPOA (Multi-Protocol Over ATM) design.
E-mail: `jhalpern@us.newbridge.com`

JOHN DRAKE is a Principal Engineer at FORE Systems, and is responsible for the development of FORE's PNNI and MPOA products. Prior to joining FORE Systems in 1995, he was a Member of Senior Technical Staff at IBM, as well as the lead designer for IBM's Broadband Network Services and Advanced Peer-Peer Networking. He also represented IBM in the PNNI subworking group. John received a B.S. in Physics and Mathematics from C-MU in 1975. He is a Senior Member of the IEEE, and has eighteen patents in the area of broadband networking. John is FORE Systems' Primary Representative to the ATM Forum and represents FORE Systems in the PNNI and MPOA sub-working groups.
E-mail: `drake@fore.com`

HALDON J. SANDICK is an advisory programmer in IBM's Networking Hardware Division in Research Triangle Park, NC. He has worked on a variety projects including the integration of APPN into SNA VTAM. The last few years he has focused on the architecture required to support various access technologies over ATM networks. He participates in several IETF Working Groups and is a contributing member of the ATM Forum's Private Network-to-Network Interface (PNNI) Sub Working Group.
E-mail: `sandick@vnet.ibm.com`

JASON JEFFORDS holds an BSEE and an MSEE from the University of New Hampshire. Since 1992 he has worked for Cabletron Systems Inc. on a wide range of networking projects including the Spectrum Enterprise Network Management platform, SecureFast Virtual Network Services (SFVNS), PNNI Routing, and various projects concerned with fault tolerant distributed systems, switched network management and control, network directory services, and policy based network management and routing. He is Cabletron's principle representative at the ATM Forum, has been an active member of the ATM Forum since 1993, and is a member of the IEEE. He is the current secretary for the ATM Forum's PNNI SWG.
E-mail: `jeffords@ctron.com`

[Ed.: You can learn more about PNNI Routing for ATM Networks by attending Workshop W810 on Thursday April 4th at NetWorld+Interop 96 in Las Vegas. The instructors for this workshop are Ross Callon and Joel Halpern.]

Call for Papers

The Second ACM International Conference on Mobile Computing and Networking, *MobiCom '96,* will be held November 10–12, 1996 at the Rye Hilton, Rye, New York, USA.

The wireless communication revolution is bringing fundamental changes to telecommunication and computing. Wide-area cellular systems and wireless LANs promise to make integrated networks a reality and provide fully distributed and ubiquitous mobile computing and communications, thus bringing an end to the tyranny of geography. Furthermore, services for the mobile user are maturing and are poised to change the nature and scope of communication. This conference, the second of an annual series, serves as the premier international forum addressing networks, systems, algorithms, and applications that support the symbiosis of portable computers and wireless networks.

Topics

Technical papers describing previously unpublished, original, completed, and not currently under review by another conference or journal are solicited on topics at the link layer and above. Topics will include, but are not limited to:

- Applications and computing services supporting the mobile user
- Network architectures, protocols or service algorithms to cope with mobility, limited bandwidth, or intermittent connectivity
- Design and analysis of algorithms for online and mobile environments
- Mobile network protocols
- Performance characterization of mobile/wireless networks and systems
- Network management for mobile and wireless networks
- Data management in mobile computing
- Service integration and interworking of wired and wireless networks
- Characterization of the influence of lower layers on the design and performance of higher layers
- Security, scalability and reliability issues for mobile/wireless systems
- Mobile computing
- Mobile agents
- Power management
- Wireless multimedia systems
- Satellite communication
- Location-dependent applications
- Distributed system aspects of mobile systems
- Adaptive applications interfaces suitable for mobile systems
- Architectures of wireless and mobile networks and systems
- Traffic integration for mobile applications

continued on next page

Call for Papers *(continued)*

All papers will be refereed by the program committee. Accepted papers will be published in conference proceedings. Papers of particular merit will be selected for publication in the ACM/Baltzer *Journal on Wireless Networks* and the ACM/Baltzer *Mobile Networks & Nomadic Applications Journal.*

Submissions

Paper submission will be handled electronically. Authors should e-mail a *PostScript* version of their full paper to:

`mobicom96@gucci.mirc.gatech.edu`

This e-mail address will become operational on March 1, 1996. In order to print the PS versions of the papers, authors should ensure that their papers meet these restrictions:

- *PostScript* version 2 or later
- No longer than 15 pages
- Fits properly on "US Letter" size paper (8.5 × 11 inches)
- Reference only Computer Modern or standard Adobe fonts (i.e., Courier, Times Roman, or Helvetica); other fonts may be used but must be included in the *PostScript* file

In addition, authors should separately e-mail the title, author names, full address and abstract of their paper to the program chairs.

All submitted papers will be judged based on their quality through double-blind reviewing where the identities of the authors are withheld from the reviewers. Authors' names should not appear on the paper or in the *PostScript* file.

Tutorials

Proposals for tutorials are solicited. Evaluation of the proposals will be based on expertise and experience of instructors, and the relevance of the subject matter. Potential instructors are requested to submit at most 5 pages, including a biographical sketch to Arvind Krishna (`krishna@watson.ibm.com`).

Panels

Panels are solicited that examine innovative, controversial, or otherwise provocative issues of interest. Panel proposals should not exceed more than 3 pages, including biographical sketches of the panelist. Potential panel organizers should contact Tom LaPorta via e-mail (`tlp@boole.att.com`).

Student participation

Papers with a student as a primary author will enter a student paper award competition. The student will receive a cash award of $500,- US Dollars. A cover letter must identify the paper as a candidate for the student paper competition.

Important dates

Submissions due:	May 1, 1996
Notification of acceptance:	July 15, 1996
Camera-ready version due:	August 31, 1996

More information

Please contact Ian F. Akyildiz (`ian@ee.gatech.edu`) or Zygmunt J. Haas (`zjh1@cornell.edu`), the Program Co-Chairs for more information.

This CFP and other ACM related activities may be found in:

`http://info.acm.org/sigcomm/mobicom96`

Future NetWorld+Interop Dates and Locations

NetWorld+Interop 96	Las Vegas, NV	April 1–5, 1996
NetWorld+Interop 96	Frankfurt, Germany	June 10–14, 1996
NetWorld+Interop 96	Tokyo, Japan	July 22–26, 1996
NetWorld+Interop 96	Atlanta, GA	September 16–20, 1996
NetWorld+Interop 96	Paris, France	October 7–11, 1996
NetWorld+Interop 96	London, England	Oct. 28–Nov. 1, 1996
NetWorld+Interop 96	Sydney, Australia	November 25–29, 1996
NetWorld+Interop 97	Singapore	April 7–11, 1997

All dates are subject to change.

More information

Call 1-800-INTEROP or +1-415-578-6900 for more information. Or send e-mail to `info@interop.com` or fax to +1-415-525-0194. For the latest information about NetWorld+Interop including *N+I Online!* as well as other SOFTBANK produced events, check our home page at `http://www.interop.com`

NetWorld+Interop is produced by SOFTBANK Exposition and Conference Company, 303 Vintage Park Drive, Foster City, California 94404–1138, USA.

Write to *ConneXions*!

We'd love to hear your comments, suggestions and questions about anything you read in *ConneXions*. Our editorial address is given below. Use it for letters to the Editor, requests for the index of back issues, questions about particular articles etc.:

ConneXions—The Interoperability Report
303 Vintage Park Drive
Suite 201
Foster City
California 94404–1138
USA
Phone: +1 415-578-6900 or 1-800-INTEROP (Toll-free in the USA)
Fax: +1 415-525-0194
E-mail: `connexions@interop.com`
URL: `http://www.interop.com`

Subscription information

For questions about your subscription please call our customer service hotline: 1-800-575-5717 or +1 610-892-1959 outside the USA. This is the number for our new subscription agency, Seybold Publications. Their fax number is +1 610-565-1858. The mailing address for subscription payments is: P.O. Box 976, Media, PA 19063–0976.

CONNEXIONS
303 Vintage Park Drive
Suite 201
Foster City, CA 94404-1138
Phone: 415-578-6900
FAX: 415-525-0194

ADDRESS CORRECTION
REQUESTED

FIRST CLASS MAIL
U.S. POSTAGE
PAID
SAN JOSE, CA
PERMIT NO. 1